Study of the B $\rightarrow$K* $\mu^+\mu$ decay with the LHCb experiment : angular analysis and measurement of the ratio RK

Samuel Coquereau



Université Pierre et Marie Curie

Ecole doctorale des Sciences de la Terre et de l'environnement et Physique de l'Univers, Paris (ED 560)

Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE)

Study of the $B^0 \to K^{*0}\mu^+\mu^-$ decay with the LHCb experiment : angular analysis and measurement of the ratio R_{K^*}

Etude de la désintégration $B^0 \to K^{*0}\mu^+\mu^-$ avec l'expérience LHCb: analyse angulaire et mesure du rapport R_{K^*}

Par Samuel Coquereau

Thèse de doctorat de physique des particules

Dirigée par Francesco Polci et Eli Ben-Haïm

Soutenue le 23 septembre 2015 devant le jury composé de :

M.	Eli	BEN-HAIM	*Directeur de thèse* (invité)
M.	Adrian	BEVAN	*Examinateur*
M.	Sébastien	DESCOTES-GENON	*Rapporteur*
M.	Patrick	KOPPENBURG	*Examinateur*
M.	Francesco	POLCI	*Co-encadrant de thèse* (invité)
M.	Jean-Paul	TAVERNET	*Président du jury*
M.	John	WALSH	*Rapporteur*

Contents

Introduction

The unprecedented large quantity of data collected by the LHCb experiment at the Large Hadron Collider LHC at CERN allows for the first time to study in details the rare decay $B^0 \to K^{*0}\mu^+\mu^-$. This decay has a strong potential to unveil the presence of new physics beyond the Standard Model of particle physics, or at least to constrain the alternative models. Indeed, the decay proceeds, in the Standard Model, via a flavour changing neutral current, described by box diagrams and loop diagrams called « electroweak penguins ». Yet undiscovered particles could appear in the loops, modifying for example the final state particles angular distributions predicted by the Standard Model.

The LHCb collaboration performed a first measurement of the angular distributions with 1 fb^{-1} of data [1] -[2]. A good agreement with the Standard Model has been observed for all the angular observables, except one, called P_5'. This is different from the prediction by 3.7 standard deviations in one q^2 bin. Several interpretations of the discrepancy have been proposed, some of them involving the presence of a new interaction mediated by a new vector boson Z'. Therefore, the analysis is currently, and for the next years with more data to come, at the core of the LHCb physics program and of the theoretical discussions.

This thesis presents in the first part the most recent angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decays performed by LHCb, using the whole dataset collected during the first run of the LHC, corresponding to an integrated luminosity of 3 fb^{-1}. The analysis is presented in all its aspects. From an extensive description of the differential decay rate and the observables of the $B^0 \to K^{*0}\mu^+\mu^-$ decay that can be seen in Sec. 1.3. A brief description of the LHCb experiment and its different sub-detector is presented in Chap. 2. The new strategy for selecting the signal events, based on a improvement on the vetoes of some background from exclusive decays of b hadrons and a significant improvement on the multivariate selection used in the combinatorial background suppression is described in Chap. 3. Then the angular analysis strategy, with a description of the maximum likelihood fit and the modeling of signal and background angular and invariant mass distribution, is detailed in Chap. 4. The systematic uncertainties and the results of this angular analysis are presented in Chap. 5, including an overview of their first theoretical interpretation. The main focus, though, is put on those parts of the analysis in which I was mainly involved, namely the combinatorial background rejection via a multivariate selection, the evaluation of the impact of low mass di-muon resonances, the parametrisation of the invariant mass distribution of the B candidates used in the likelihood fit, the study of the background angular distributions.

The second part of the thesis presents another analysis based on the $B^0 \to K^{*0}\mu^+\mu^-$ decay: the status of the measurement of the ratio R_{K^*}. This measurement is a test of the lepton flavour

[1] LHCb collaboration, R. Aaij et al., Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$

[2] LHCb collaboration, R. Aaij et al., Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0}\mu^+\mu^-$

universality, which is of particular interest nowadays, after that some experimental results have shown hints of a possible violation of it. This is the case, for example, of the recent observation of a 2.6 standard-deviations difference from unity of the ratio $R_K = \mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+e^+e^-)$ [3]. It is worth to note that some theoretical approaches relate the deviation in P_5' to the one in the R_K measurement, as being to manifestations of a common source of new physics[4]-[5]

The measurement of R_{K^*}, confirming or not the R_K deviation, will be shedding more light on this puzzle. It will be performed in three regions of q^2, that are sensitive to different type of physics. Since the analysis is currently still ongoing, no results will be shown in this thesis. Nevertheless, the main analysis procedure will be described in Chap. 6, with a focus on the low q^2 region and on the aspects on which I am directly involved.

[3] LHCb collaboration, R. Aaij et al., Test of Lepton Universality Using $B^+ \to K^+\ell^+\ell^-$ Decays,

[4] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, Quark flavor transitions in $L_\mu - L_\tau$ models

[5] A. Crivellin, G. D'Ambrosio, and J. Heeck, Explaining $h \to \mu^\pm\tau^\mp$, $B^0 \to K^{*0}\mu^+\mu^-$ and $B \to K\mu^+\mu^-/B \to Ke^+e^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$

Chapter 1

Theoretical contest of particle physics

Sommaire

This chapter describes the theoretical motivation for studying the electroweak penguin decays and more specifically the $B^0 \to K^{*0}\mu^+\mu^-$ decay. A brief description of the Standard Model will be given in Section 1.1. In the Section 1.2, a description of the Flavour Changing Neutral Currents (FCNC) will be given. Finally a description of the $B^0 \to K^{*0}\mu^+\mu^-$ decay can be seen in Section. 1.3

1.1 Standard Model

The Standard Model (SM) of particle physics is based on a quantum field theory and is described by a $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry. The $SU(3)_C$ describes the quarks

and the strong interactions, while $SU(2)_L \times U(1)_Y$ unifies the electromagnetism and the weak interaction. Additional fields, representing the bosons, appear since these groups have to be locally gauge invariant. This requirement implies massless boson and fermion fields.

The dynamics of the SM particles can be expressed by the following Lagrangian:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{Higgs}}, \tag{1.1}$$

where $\mathcal{L}_{\text{EW}}$ describes the electroweak sector, $\mathcal{L}_{\text{QCD}}$ the flavour sector and the final component is the Higgs mechanism responsible for the mass of the fermions.

The particles contained in the Standard Model are summarised in Fig. 1.1. The fermionic matter particles are split into two categories: the quarks, sensible to all interactions and the leptons, which are not sensible to the strong interaction. These two categories are arranged in three generations of increasing mass. For the leptons, each generation contains a massive particle, (the electron, the muon or the tau) and a corresponding massless neutrino[1]. We also have in addition to the fermions, the gauge bosons, vector of the different interactions and the Higgs boson which gives masses to the particles.

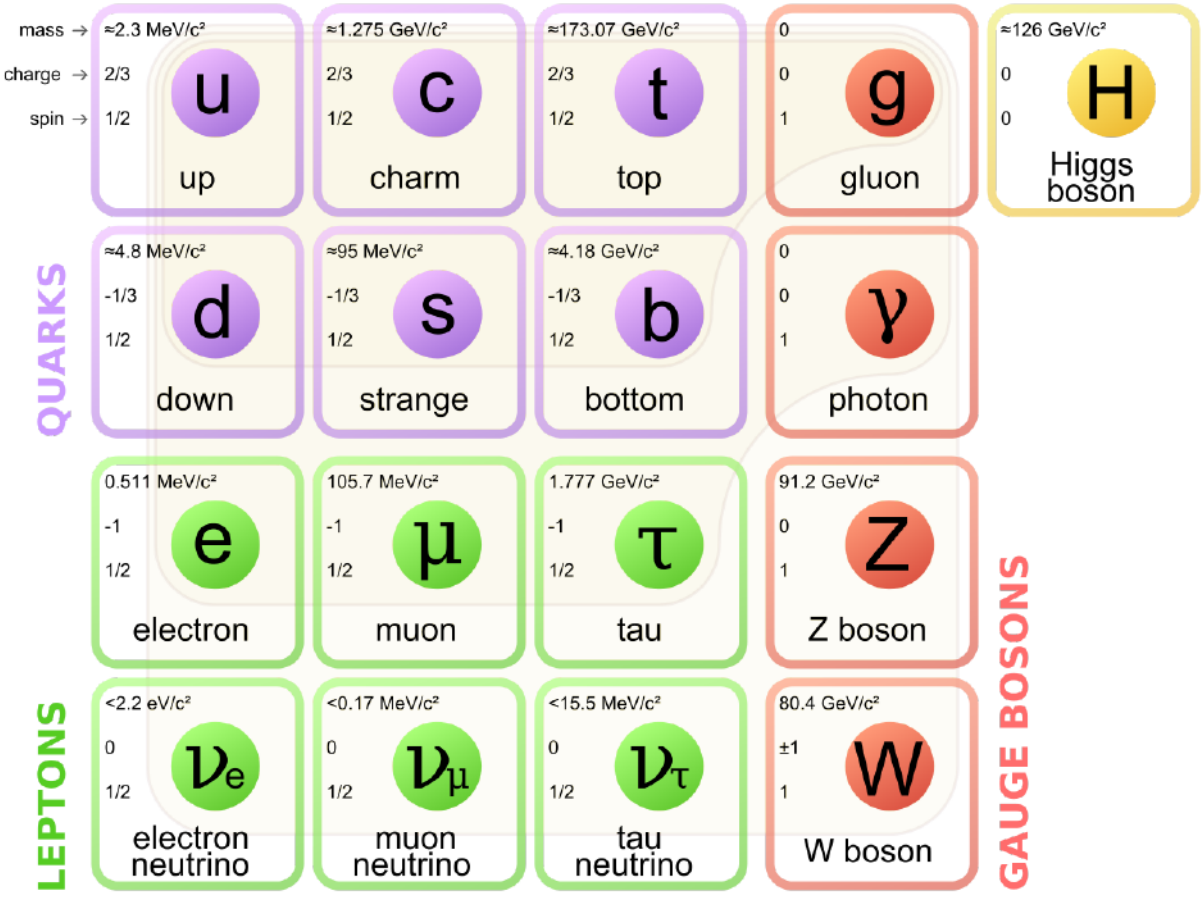


Figure 1.1: List of Standard Model particles.

[1]The observation of neutrino oscillation have implied that neutrinos have masses. Extension to the Standard Models, as the See-Saw mechanism, are needed to obtained massive neutrinos

1.1.1 Electromagnetism

The electromagnetic interaction is described in the quantum field theory by the *Quantum ElectroDynamics* (QED).

In quantum field theory a fermion field ψ, of mass m, is obeying to the Dirac equation:

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \tag{1.2}$$

where γ^μ corresponding to Dirac matrices. Leading to the following Lagrangian:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi, \qquad \text{with } \bar{\psi} = \psi^\dagger \gamma^0 \tag{1.3}$$

The Lagrangian is invariant under global gauge transformation. But in quantum field theory, interactions are introduced via a local gauge transformation keeping the physics invariant. In the case of the U(1) symmetry, the gauge transformation is defined as :

$$\begin{aligned} \psi(x) &\to \psi'(x) = e^{i\alpha(x)}\psi(x) \\ \bar{\psi}(x) &\to \bar{\psi}'(x) = e^{-i\alpha(x)}\bar{\psi}(x) \end{aligned} \tag{1.4}$$

where $\alpha(x)$ is an arbitrary function describing the local transformation. The Lagrangian (1.3) is not invariant under this transformation. Therefore a covariant derivative and a vector field A_μ of spin 1, are introduced to restore the invariance:

$$D_\mu = \partial_\mu - ieA_\mu \tag{1.5}$$

where A_μ is the field representing the photon.

This vector field transforms under the $U(1)$ as:

$$A_\mu \to A_\mu + \frac{1}{e}\partial_\mu \alpha \tag{1.6}$$

The dynamics of this electromagnetic gauge boson, are given by the Lagrangian:

$$\mathcal{L} = -\frac{1}{2}F_{\mu\nu}F^{\mu\nu} \tag{1.7}$$

where $F_{\mu\nu}$ is the Faraday tensor defined as:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.8}$$

Finally, the complete QED Lagrangian is described by:

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu \psi A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{1.9}$$

It's important to notice that a mass term for the photon, of the form $m_\gamma^2 A_\mu A^\mu$, would break the local gauge invariance. As a consequence, the $U(1)$ symmetry implies that the photon is massless.

1.1.2 Weak interaction

The weak interaction is described by a gauge theory based on the $SU(2)_L$ symmetry, where L indicates that only left handed fermions are sensitive to the the weak interaction. Left and

right-handed fermions are indicated by the ψ_L and ψ_R fields respectively. To conserve this symmetry, a new quantum number is defined: the weak isospin, I. The left-handed fermions are represented in SU(2) as doublets of weak isospin $I = \frac{1}{2}$:

$$\begin{aligned} \ell_L &= \begin{pmatrix} e_L \\ \nu_{e,L} \end{pmatrix}, \begin{pmatrix} \mu_L \\ \nu_{\mu,L} \end{pmatrix}, \begin{pmatrix} \tau_L \\ \nu_{\tau,L} \end{pmatrix}, \\ q_L &= \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}, \end{aligned} \tag{1.10}$$

While the right-handed fermions, un-sensitive to the weak interaction, are represented by SU(2) singlets with a null weak isospin ($I = 0$):

$$\begin{aligned} \ell_R &= e_R, \mu_R, \tau_R, \nu_{(e,R)}, \nu_{(\mu,R)}, \nu_{(\tau,R)}, \\ q_R &= u_R, d_R, c_R, s_R, b_R, t_R \end{aligned} \tag{1.11}$$

The $SU(2)_L$ symmetry transforms the left-handed doublets as:

$$\psi_L \rightarrow e^{i\theta_i \frac{\sigma_i}{2}} \psi_L \tag{1.12}$$

where σ_i (i=1,2,3) are the Pauli matrices and θ_i is function of the position. To keep the Lagrangian locally invariant, a covariant derivative is build and defined as:

$$D_\mu = \partial_\mu - ig\frac{\sigma_i}{2} W_\mu^i \tag{1.13}$$

where W_μ^i are the three gauge fields of SU(2) and g the weak coupling constant. These W_μ^i are transformed under $SU(2)$ as:

$$W_\mu^i \rightarrow W_\mu^i - \frac{1}{g}\partial_\mu \theta^i + \epsilon^{ijk}\theta^j G_\mu^k \tag{1.14}$$

where ϵ^{ijk} are the structure constants of the $SU(2)_L$ group. The dynamics of these fields are described by the $-\frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu}$term, with $W_{\mu\nu}^i$ defined as:

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk} W_\mu^j W_\nu^k \tag{1.15}$$

Finally, the complete weak interaction Lagrangian is described by:

$$\mathcal{L}_W = \bar{\psi}(i\gamma^\mu \partial_\mu)\psi + g\bar{\psi}\gamma^\mu \sigma_i \psi W_\mu^i - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu}. \tag{1.16}$$

The mass term $m\bar{\psi}\psi$ is not present since it should be written as $m\bar{\psi}\psi = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)$, which is not invariant under the $SU(2)_L$ symmetry. This implies that fermions should be massless. The Higgs mechanism described in Section 1.1.5 will introduce the mass of the fermions.

1.1.3 Electroweak sector

In the 1960s the electromagnetism and the weak interaction have been unified, by Glashow, Weinberg and Salam, under the same theory: the electroweak theory. The result is a group of symmetry $SU(2)_L \times U(1)_Y$, were the $SU(2)_L$ is the symmetry group for the weak isospin I,

described in Section 1.1.2, and Y is the weak hypercharge defined as :

$$Y = 2(Q - I_3) \tag{1.17}$$

where Q is the electric charge and $I_3 = \pm 1/2$ is a projection of the weak isospin.

In a similar way as the QED introduced the vector field A_μ to be gauge invariant, the gauge invariance associated to $U(1)_Y$ introduces the vector field B_μ, while the gauge invariance associated to $SU(2)_L$ implies the presence of a triplet of vector fields W^i_μ, described in the Section 1.1.2. The Lagrangian invariance under the $SU(2)_L \times U(1)_Y$ group is insured by the creation of two covariant derivatives:

$$D_{\mu,L} = \partial_\mu + ig\frac{\sigma_i}{2}W^i_\mu + ig'\frac{Y}{2}B_\mu \tag{1.18}$$

$$D_{\mu,R} = \partial_\mu + ig'\frac{Y}{2}B_\mu \tag{1.19}$$

where g and g' are the coupling constants of the $SU(2)_L$ and $U(1)_Y$ symmetries respectively.

Finally, the complete electroweak Lagrangian is described by:

$$\mathcal{L}_{EW} = i\bar{\psi_L}\gamma^\mu D_{\mu,L}\psi_L + i\bar{\psi_R}\gamma^\mu D_{\mu,R}\psi_R - \frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.20}$$

In this representation, the γ, Z^0 and $W^\pm$ bosons are described by the following fields:

$$A_\mu = B_\mu \cos\theta_W + W^3_\mu \sin\theta_W \tag{1.21}$$

$$Z_\mu = -B_\mu \sin\theta_W + W^3_\mu \cos\theta_W \tag{1.22}$$

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \tag{1.23}$$

where θ_W is the Weinberg or weak mixing angle.

The electromagnetic coupling constant α_e, can be derived by the following equation:

$$\alpha_e = g \sin\theta_W = g' \cos\theta_W \tag{1.24}$$

The Weinberg angle could also be expressed with the coupling constants g and g' by:

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \qquad \text{and} \qquad \sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \tag{1.25}$$

1.1.4 Strong interaction

The strong interaction is described in the quantum field theory by the *Quantum Chromo-Dynamics* (QCD). This interaction acts between quarks and gluons. The QCD has been build on the base of the QED by Gell-Mann and Zweig in the 1960s. This theory is based on gauge invariance associated to the SU(3) gauge group and each of the quarks form a colour triplets,

$$q = \begin{pmatrix} q_R \\ q_G \\ q_B \end{pmatrix}, \tag{1.26}$$

where R,G and B are the three colour indices for red, green and blue respectively.

To insure the gauge invariance, a covariant derivative is defined as:

$$D_\mu = \partial_\mu - ig_s \frac{\lambda_\alpha}{2} G_\mu^\alpha \tag{1.27}$$

where λ_α are the Gell-Mann matrices and G_μ^α are the eight gauge fields corresponding to the gluons. These fields transform under $SU(3)$ as:

$$G_\mu^\alpha \to G_\mu^\alpha - \frac{1}{g_s}\partial_\mu \theta^\alpha + f^{\alpha\beta\gamma}\theta^\beta G_\mu^\gamma \tag{1.28}$$

where $f^{\alpha\beta\gamma}$ are the structure constants of the SU(3) group.

The dynamics of the gluon field is described by the tensors:

$$G_{\mu\nu}^\alpha = \partial_\mu G_\nu^\alpha - \partial_\nu G_\mu^\alpha + g_s f^{\alpha\beta\gamma} G_\mu^\beta G_\nu^\gamma \tag{1.29}$$

Since $SU(3)$ is a non-commutative group, the kinetic term has a component multiplying two G_μ, describing the self-coupling between the gluons.

Finally the QCD Lagrangian is given by:

$$\mathcal{L}_{QCD} = \sum_f \left(\bar{q_f}(i\gamma^\mu \partial_\mu - m) q_f + g_s \bar{q_f} \gamma^\mu \lambda_\alpha q_f G_\mu^\alpha \right) - \frac{1}{4} G_{\mu\nu}^\alpha G_\alpha^{\mu\nu} \tag{1.30}$$

where f represent the quarks flavour.

The mass term for the gluon breaks the SU(3) gauge invariance, which explains why the gluons are massless.

1.1.5 The Higgs mechanism

The gauge invariance of the $SU(2)_L \times U(1)_Y$ leads to the fact that the four gauge bosons, as well as the fermions, are massless. Nevertheless we know experimentally that the $W^\pm$ and Z^0 bosons, the quarks and the charged leptons are massive.

The Higgs mechanism is a way to solve this problem. It generates the particles masses dynamically, through spontaneous symmetry breaking of the $SU(2)_L \times U(1)_Y$ symmetry.

Starting with a $SU(2)_L$ doublet of complex scalar defined as:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.31}$$

we can define a Lagrangian which preserves the $SU(2)_L \times U(1)_Y$ symmetry:

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D^\mu \phi) - (\lambda(\phi^\dagger \phi)^2 + \mu^2 \phi^\dagger \phi), \tag{1.32}$$

where $-(D_\mu\phi)^\dagger(D^\mu\phi)$ is the kinetic term for a scalar particle, $\lambda > 0$ the coupling of a four particle vertex and μ represents the mass of the field for $\mu^2 > 0$. In the case where $\mu^2 < 0$, the potential $V(\phi) = \lambda(\phi^\dagger\phi)^2 + \mu^2\phi^\dagger\phi$ reaches its minimum for ϕ values of:

$$|\phi| = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{1}{\sqrt{2}}\nu, \tag{1.33}$$

where ν is the vacuum expectation value, measured to be $\nu = 246\,\text{GeV}$. This represents a degenerate ground state. We can choose one of the minima as the vacuum without loss of generality, since any other minima could be reached by a gauge transformation:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix}, \tag{1.34}$$

where $h(x)$ is a perturbative expansion around the minimum value, a real scalar field which corresponds to a spin 0 boson: the Higgs boson.

The mechanism by which this particular field is chosen from the degeneration of the ground state is called spontaneous symmetry breaking.

Developing the kinetic term of the Equation. (1.36), $(D_\mu\phi)^\dagger(D^\mu\phi)$, we can obtained the mass terms of the gauge bosons:

$$\begin{aligned} m_A &= 0 \\ m_W &= \frac{gv}{2} = m_Z \cos\theta_W \\ m_Z &= \frac{v}{2}\sqrt{g^2 + g'^2} \end{aligned} \tag{1.35}$$

Introducing in the Lagrangian the Yukawa couplings between fermions and the Higgs field, a mass term for the fermions appears. As example the Lagrangian of the electron is defined as:

$$\begin{aligned} \mathcal{L} &= -\lambda_e(\bar{\psi}_L\phi\psi_R + \bar{\psi}_R\phi^\dagger\psi_L) \\ \mathcal{L} &= -\frac{\lambda_e}{\sqrt{2}}\left[(\bar{\nu}_e, \bar{e})_L\begin{pmatrix} 0 \\ v+h \end{pmatrix}e_R + \bar{e}_R(0, v+h)\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L\right] \\ \mathcal{L} &= -\frac{\lambda_e v}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{\lambda_e}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L)H, \\ \mathcal{L} &= -m_e\bar{e}e - \frac{m_e}{v}\bar{e}eH, \end{aligned} \tag{1.36}$$

where λ_e is the Yukawa coupling of the electron. The first term of the Lagrangian gives the mass to the electron $m_e = \frac{\lambda_e v}{\sqrt{2}}$ while the second term defines the coupling, $\frac{m_e}{v}$, between the Higgs field and the electron. Note that the Higgs couples to fermions proportionally to their masses.

The mass of the Higgs boson is given by:

$$m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda\nu^2} \tag{1.37}$$

The λ parameter is free in the SM, so the mass of the Higgs boson is not predicted by it. The 4th of July 2012 the ATLAS and CMS collaborations have announced the detection of a new boson with a mass of around 125 GeV/c^2. In March 2013, the measurement of the total angular momentum and the parity of this new particle have been presented: they are in agreement with those predicted for the Higgs boson: $J^P = 0^+$.

1.1.6 The CKM mechanism

The description of the electroweak interaction presented in Section 1.1.3, doesn't explain the couplings between the different quarks generations. This mixing is taken into account by rewriting the charged current in the weak interaction Lagrangian as:

$$-\frac{g}{2\sqrt{2}}(\bar{u}'^i_L\gamma^\mu V_{ij}d^j_L W^+_\mu + \bar{d}'^j_L\gamma^\mu V_{ij}u^i_L W^-_\mu) \tag{1.38}$$

where $V_{ij} = (U_u^\dagger U_d)_{ij}$ corresponds to the element of the Cabibbo-Kobayashi-Maskawa (CKM) matrix [1, 2]:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.39}$$

This 3×3 matrix represents the change of basis between the quark mass eigenstates and the quark interaction eigenstates. It matrix can be parametrised by three angles and one complex phase.

$$V = \begin{pmatrix} c_{12} & s_{12}e^{i\delta} & 0 \\ -s_{12}e^{-i\delta} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23}e^{i\delta} \\ 0 & -s_{23}e^{-i\delta} & c_{23} \end{pmatrix} \times \begin{pmatrix} c_{13} & 0 & s_{13}e^{i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \tag{1.40}$$

where, $i, j = 1, 2, 3$ with $i \neq j$, $c_{ij} = \cos(\theta_{ij})$ and $s_{ij} = \sin(\theta_{ij})$. The δ is the irreducible complex phase and the θ_{ij} are the three angles of the parametrisation. The θ_{12}, responsible of the u-s quark mixing, is known as the Cabibbo angle and measured to be $s_{12} = 0.22$.

The CKM matrix can also be written in terms of the Wolfenstein parametrisation [3]:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.41}$$

where A, λ and ρ are the real parameters and η the imaginary part which is responsible for the only source of CP violation in the SM.

1.1.7 Success of the Standard Model and its limits

Considering, the discovery of the Z^0 and $W^\pm$ bosons and of the twelve fermions, the observation of the CP violation and recently the Higgs boson discovery, the Standard Model has been a very predictive theory. But despite its success, it has also some limits:

- The Standard Model has 19 free parameters, namely the 9 fermions masses, the 3 coupling constants, the Z^0 and Higgs Boson masses, the 4 parameters from the CKM matrix and the strong phase describing the CP violation in the strong sector. This large number of parameters are not predicted and have to be measured experimentally.

- The Standard Model has no explanation for the values of the masses and their huge difference between the families, neither has an explanation for the number of families observed.

- Neutrino are massive particles, as demonstrated by the observation of their oscillation, while no mechanism can give masses to neutrinos in the Standard Model.

- The measured mass of the Higgs boson implies, a huge cancelation between the mass and the quantum corrections. This issue leads to a motivation for a New Physics at the TeV scale, which will provide a natural correction of the fine-tuning problem.

- Gravity is not described by the Standard Model.

- The observed matter, described by the model, represent only 4.8 % of the total mass of the universe.
- The amount of CP violation is ten orders of magnitude lower that what is needed to produce the matter-antimatter asymmetry of the universe after the Big Bang.

All these reasons lead many physicists to think that the Standard Model is a part of a more global model, but, for the moment no manifestation of this "New Physics" has been found. The study of the $B^0 \to K^{*0}\mu^+\mu^-$ and the measurement of R_{K^*}, described in this thesis, are a way to search for manifestation of beyond Standard Model phenomenon.

1.2 Flavour changing neutral currents

Flavour-changing neutral currents (FCNCs) are processes involving a quark transition between two up-type or two down-type quarks, *i.e.* a change of quarks families without a change in electric charge. These FCNC processes are predicted to be rare within the SM, since they occur only at loop level where the mediating $W^{\pm}$ bosons are virtual. Measurements of FCNCs processes are model-independent tests of any new physics model which modifies the properties of the decay. Indeed, new particles could appear as virtual particles in the loops, causing such modifications respect to the Standard Model predictions.

The $b \to s$ transitions are an example of FCNC decays, where a photon or dilepton is produced within a box or a penguin diagram as shown on Figure. 1.2

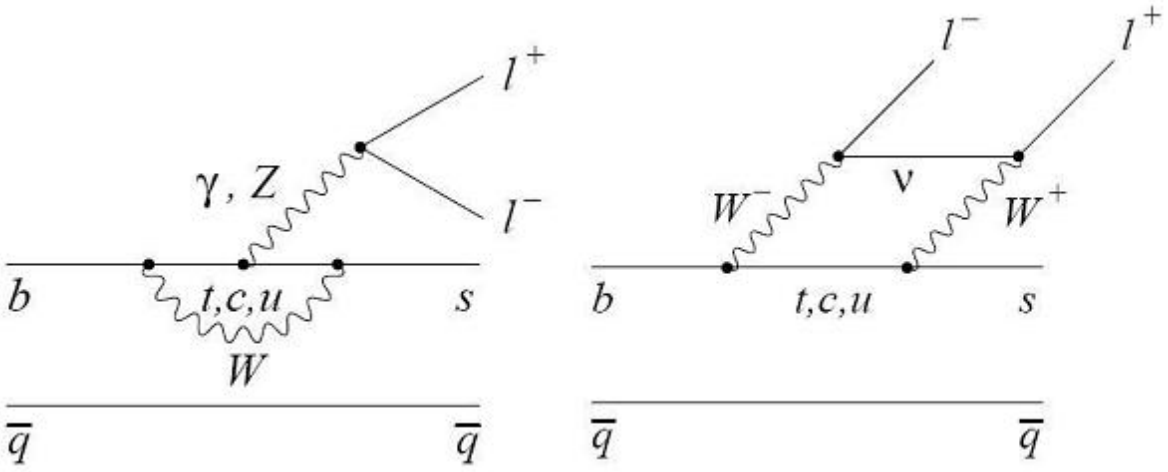


Figure 1.2: Lowest order penguin (left) and box (right) SM diagrams of a $b \to s\ell^+\ell^-$ transition.

1.2.1 Effective field theory in flavour physics

The formalism of b decays can be expressed with the Operator Product Expansion (OPE), which integrates out all degrees of freedom above some energy scale μ. This approach is valid if μ is larger than the relevant energy scale of the decay, which for the B physics is around 5 GeV/c^2. This formalism is similar to Fermi's effective theory of weak decays, in which box or penguin diagram is interpreted as a four point interaction.

The matrix element for a high energy process of an effective hamiltonian is defined as:

$$< f|H_{\text{eff}}|i >= \sum_j C_j(\mu_s) < f|\mathcal{O}(\mu_s)_j|i > \tag{1.42}$$

where the $\mathcal{C}_j$, named Wilson coefficients, are complex numbers which encapsulate the high energy, short distance contribution, while the $\mathcal{O}_i$ are local operators to particular gauge structure which encode the low energy contribution from quarks. The scale μ_s is the renormalisation scale where the operators and the Wilson coefficients are evaluated. The advantage of this formalism is that Wilson coefficients are independent of the underlying process and can include contributions from new physics models.

The effective Hamiltonian for a $b \to s\ell^+\ell^-$ decay [4] is:

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \frac{\alpha_e}{4\pi} \sum_i C_i(\mu_s)\mathcal{O}(\mu_s) \tag{1.43}$$

where G_F is the Fermi constant, V_{ij} are CKM matrix elements, α_e the fine structure constant.

The local operators important for these b electroweak penguin decays are :

$$\begin{aligned}
\mathcal{O}_7 &= \frac{m_b}{e}\bar{s}\sigma^{\mu\nu}P_R b F_{\mu\nu}, & \mathcal{O}'_7 &= \frac{m_b}{e}\bar{s}\sigma^{\mu\nu}P_L b F_{\mu\nu},\\
\mathcal{O}_9 &= \bar{s}\gamma_\mu P_L b\bar{\ell}\gamma^\mu\ell, & \mathcal{O}'_9 &= \bar{s}\gamma_\mu P_R b\bar{\ell}\gamma^\mu\ell,\\
\mathcal{O}_{10} &= \bar{s}\gamma_\mu P_L b\bar{\ell}\gamma^\mu\gamma_5\ell, & \mathcal{O}'_{10} &= \bar{s}\gamma_\mu P_R b\bar{\ell}\gamma^\mu\gamma_5\ell,
\end{aligned}$$

where $P_{L,R} = (1 \mp \gamma_5)/2$ are a left/right handed chiral projection.

The $\mathcal{O}_7$ is the electromagnetic operator, corresponding to the emission of a photon. The $\mathcal{O}_9$ and $\mathcal{O}_{10}$ are operators describing the Z penguin and W box diagrams. The $\mathcal{O}'$ operators refer to right-handed couplings, obtained by replacing $P_L \leftrightarrow P_R$ in the $\mathcal{O}_i$ operators and conducts, in the SM, to a m_s/m_b suppression for the Wilson coefficients C'_i with respect to the C_i.

1.3 The $B^0 \to K^{*0}\mu^+\mu^-$ decay

The $B^0 \to K^*(\to K\pi)\mu^+\mu^-$ decay is a $b \to s$ transition mediated by a flavour changing neutral current, and the relative Feynmann diagrams are shown in Figure 1.2 at leading order. It is a particularly interesting decay due to the large number of Wilson coefficients to which it is sensitive. Indeed, while decays such as $B \to K^*\gamma$ or $B_s \to \mu^+\mu^-$ [2] can access to one Wilson coefficient only , $C_7^{(\prime)}$ or $C_{10}^{(\prime)}$ respectively, the $B^0 \to K^{*0}\mu^+\mu^-$ decay allows to measure $C_7^{(\prime)}$ in the very low-q^2 region, a mixture of $C_7^{(\prime)} - C_9^{(\prime)}$ in the low-q^2 region and $C_9^{(\prime)}$ and $C_{10}^{(\prime)}$ dominating the high-q^2 region. Moreover, from an experimental point of view, channels with muons are easier to trigger on at LHCb, which makes easier to reconstruct this decay respect to $B^0 \to K^*(\to K\pi)e^+e^-$, the corresponding decay with electrons in the final state instead of muons. This provides a larger number of events available to measure a multiplicity of observables: total and differential branching fraction as function of q^2, CP asymmetries [5], isospin asymmetries [6], and especially many observables related to the angular distributions of the four particles in the final-state. Theoretical predictions in the Standard Model framework are available for all these

[2] The $B_s \to \mu^+\mu^-$ decay can also test non-Standard Model Wilson coefficient $C_{S,P}^{(\prime)}$

observables. The experimental measurements can be compared with the predictions to search for hints of new physics.

The measurement of these angular observables is the main topic of the first part of this thesis and they will be described in this section.

1.3.1 The angular basis

Different definitions of the decay angles exist in theory and experiment. Details on the differences between the convention chosen by a majority of the theory publications [4, 7] and the convention adopted by LHCb for the analysis presented in this thesis and the previous publications [8, 9] are given in Ref. [10].

The angular basis used in the analysis is illustrated in Fig. 1.3.

- The angle θ_ℓ is defined as the angle between the direction of the μ^+ (μ^-) in the dimuon rest frame and the direction of the dimuon in the B^0 ($\overline{B}^0$) rest frame, explicitly for the B^0:

$$\cos\theta_\ell = \left(\hat{p}_{\mu^+}^{(\mu^+\mu^-)}\right)\cdot\left(\hat{p}_{\mu^+\mu^-}^{(B^0)}\right) = \left(\hat{p}_{\mu^+}^{(\mu^+\mu^-)}\right)\cdot\left(-\hat{p}_{B^0}^{(\mu^+\mu^-)}\right) , \tag{1.44}$$

  and for the $\overline{B}^0$ decay:

$$\cos\theta_\ell = \left(\hat{p}_{\mu^-}^{(\mu^+\mu^-)}\right)\cdot\left(\hat{p}_{\mu^+\mu^-}^{(\overline{B}^0)}\right) = \left(\hat{p}_{\mu^-}^{(\mu^+\mu^-)}\right)\cdot\left(-\hat{p}_{\overline{B}^0}^{(\mu^+\mu^-)}\right) , \tag{1.45}$$

- The angle θ_K is defined as the angle between the direction of the kaon in the K^{*0} ($\overline{K}^{*0}$) rest frame and the direction of the K^{*0} ($\overline{K}^{*0}$) in the B^0 ($\overline{B}^0$) rest frame.

$$\cos\theta_K = \left(\hat{p}_{K^+}^{(K^{*0})}\right)\cdot\left(\hat{p}_{K^{*0}}^{(B^0)}\right) = \left(\hat{p}_{K^+}^{(K^{*0})}\right)\cdot\left(-\hat{p}_{B^0}^{(K^{*0})}\right) \tag{1.46}$$

  for the B^0 and

$$\cos\theta_K = \left(\hat{p}_{K^-}^{(K^{*0})}\right)\cdot\left(\hat{p}_{K^{*0}}^{(\overline{B}^0)}\right) = \left(\hat{p}_{K^-}^{(K^{*0})}\right)\cdot\left(-\hat{p}_{\overline{B}^0}^{(K^{*0})}\right) \tag{1.47}$$

  for the $\overline{B}^0$ decay.

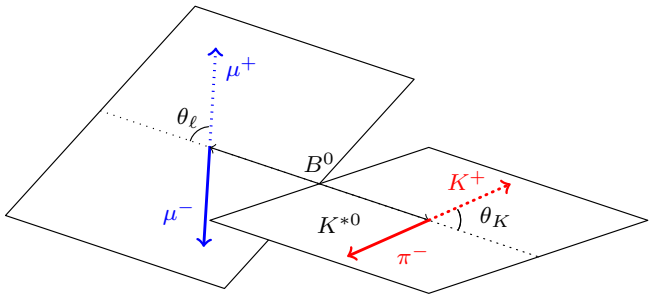
- The angle ϕ is the angle between the plane containing the μ^+ and μ^- and the plane containing the kaon and pion from the K^{*0}. The definition of the angle ϕ for the B^0 is given by:

$$\cos\phi = \left(\hat{p}_{\mu^+}^{(B^0)}\times\hat{p}_{\mu^-}^{(B^0)}\right)\cdot\left(\hat{p}_{K^+}^{(B^0)}\times\hat{p}_{\pi^-}^{(B^0)}\right) , \tag{1.48}$$

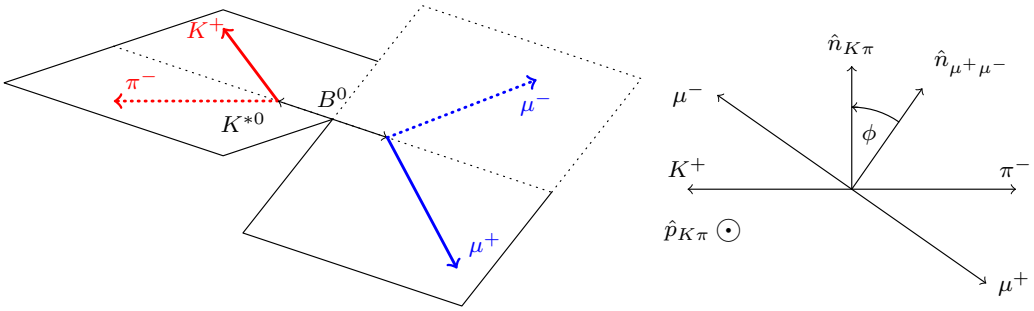
$$\sin\phi = \left[\left(\hat{p}_{\mu^+}^{(B^0)}\times\hat{p}_{\mu^-}^{(B^0)}\right)\times\left(\hat{p}_{K^+}^{(B^0)}\times\hat{p}_{\pi^-}^{(B^0)}\right)\right]\cdot\hat{p}_{K^{*0}}^{(B^0)} \tag{1.49}$$

  for the B^0 and by:

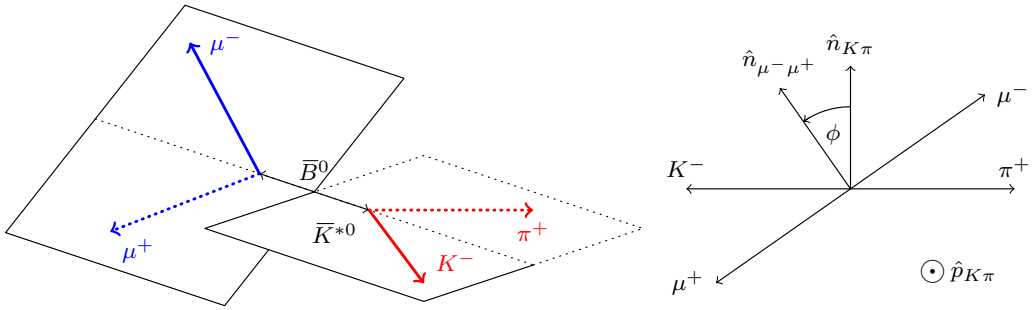
$$\cos\phi = \left(\hat{p}_{\mu^-}^{(\overline{B}^0)}\times\hat{p}_{\mu^+}^{(\overline{B}^0)}\right)\cdot\left(\hat{p}_{K^-}^{(\overline{B}^0)}\times\hat{p}_{\pi^+}^{(\overline{B}^0)}\right) , \tag{1.50}$$



(a) θ_K and θ_ℓ definitions for the B^0 decay



(b) ϕ definition for the B^0 decay



(c) ϕ definition for the $\overline{B}^0$ decay

Figure 1.3: Graphical representation of the angular basis used for $B^0 \to K^{*0}\mu^+\mu^-$ and $\overline{B}^0 \to \overline{K}^{*0}\mu^+\mu^-$ decays in this paper. The notation $\hat{n}_{ab}$ is used to represent the normal to the plane containing particles a and b in the B^0 (or $\overline{B}^0$) rest frame:

$$\sin\phi = -\left[\left(\hat{p}_{\mu^-}^{(\overline{B}^0)} \times \hat{p}_{\mu^+}^{(\overline{B}^0)}\right) \times \left(\hat{p}_{K^-}^{(\overline{B}^0)} \times \hat{p}_{\pi^+}^{(\overline{B}^0)}\right)\right] \cdot \hat{p}_{\overline{K}^{*0}}^{(\overline{B}^0)} \tag{1.51}$$

for the $\overline{B}^0$ decay.

The $\hat{p}_X^{(Y)}$ are unit vectors describing the direction of a particle X in the rest frame of the system Y. In every case the particle momenta are first boosted to the B^0 (or $\overline{B}^0$) rest frame. In this basis, the angular definition for the $\overline{B}^0$ decay is a CP transformation of that for the B^0 decay.

1.3.2 The differential decay rate

The differential decay rate of $B^0 \to K^{*0}\mu^+\mu^-$ and $\overline{B}^0 \to \overline{K}^{*0}\mu^+\mu^-$ decays, in terms of q^2 and the three angles, θ_K, θ_l and ϕ, is given by:

$$\begin{aligned}\frac{\mathrm{d}^4\Gamma[\overline{B}^0 \to \overline{K}^{*0}\mu^+\mu^-]}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2} =& \frac{9}{32\pi}\sum_i J_i(q^2) f_i(\cos\theta_\ell, \cos\theta_K, \phi) \\ =& \frac{9}{32\pi}\Big[\, J_1^s \sin^2\theta_K + J_1^c \cos^2\theta_K \; + \\ & J_2^s \sin^2\theta_K \cos 2\theta_\ell + J_2^c \cos^2\theta_K \cos 2\theta_\ell \; + \\ & J_3 \sin^2\theta_K \sin^2\theta_\ell \cos 2\phi + J_4 \sin 2\theta_K \sin 2\theta_\ell \cos\phi \; + \\ & J_5 \sin 2\theta_K \sin\theta_\ell \cos\phi + J_6^s \sin^2\theta_K \cos\theta_\ell \; + \\ & J_7 \sin 2\theta_K \sin\theta_\ell \sin\phi + J_8 \sin 2\theta_K \sin 2\theta_\ell \sin\phi \; + \\ & J_9 \sin^2\theta_K \sin^2\theta_\ell \sin 2\phi \;\Big].\end{aligned} \tag{1.52}$$

Here, the q^2 dependent angular observables $J_i(q^2)$ are given by:

$$
\begin{aligned}
J_1^s &= \frac{(2+\beta_\mu^2)}{4}\left[|A_\perp^L|^2+|A_\parallel^L|^2+|A_\perp^R|^2+|A_\parallel^R|^2\right]+\frac{4m_\mu^2}{q^2}\Re e(A_\perp^L A_\perp^{R*}+A_\parallel^L A_\parallel^{R*}) \\
J_1^c &= |A_0^L|^2+|A_0^R|^2+\frac{4m_\mu^2}{q^2}\left[|A_t|^2+2\Re e(A_0^L A_0^{R*})\right] \\
J_2^s &= \frac{\beta_\mu^2}{4}\left[|A_\perp^L|^2+|A_\parallel^L|^2+|A_\perp^R|^2+|A_\parallel^R|^2\right] \\
J_2^c &= -\beta_\mu^2\left[|A_0^L|^2+|A_0^R|^2\right] \\
J_3 &= \frac{\beta_\mu^2}{2}\left[|A_\perp^L|^2-|A_\parallel^L|^2+|A_\perp^R|^2-|A_\parallel^R|^2\right] \\
J_4 &= \frac{\beta_\mu^2}{\sqrt{2}}\left[\Re e(A_0^L A_\parallel^{L*})+\Re e(A_0^R A_\parallel^{R*})\right] \\
J_5 &= \sqrt{2}\beta_\mu\left[\Re e(A_0^L A_\perp^{L*})-\Re e(A_0^R A_\perp^{R*})\right] \\
J_6^s &= 2\beta_\mu\left[\Re e(A_\parallel^L A_\perp^{L*})-\Re e(A_\parallel^R A_\perp^{R*})\right] \\
J_7 &= \sqrt{2}\beta_\mu\left[\Im m(A_0^L A_\parallel^{L*})-\Im m(A_0^R A_\parallel^{R*})\right] \\
J_8 &= \frac{\beta_\mu^2}{\sqrt{2}}\left[\Im m(A_0^L A_\perp^{L*})+\Im m(A_0^R A_\perp^{R*})\right] \\
J_9 &= \beta_\mu^2\left[\Im m(A_\parallel^{L*} A_\perp^L)+\Im m(A_\parallel^{R*} A_\perp^R)\right]
\end{aligned}
\tag{1.53}
$$

with $\beta_\mu^2 = (1-4m(\mu)^2/q^2)$. The angular distribution therefore depends on 7 q^2 dependent complex amplitudes ($A_0^{L,R}$, $A_\parallel^{L,R}$, $A_\perp^{L,R}$ and A_t) corresponding to different polarisation states of the $B \to K^* V^*$ decay. The K^{*0} is on-shell and has three polarisation states, $\epsilon(+,-,0)$. The V^* is off-shell and has 4 possible states, $\epsilon(+,-,0,t)$. The amplitude A_t corresponds to a longitudinal polarisation of the K^{*0} and time-like polarisation of the dimuon system. It is suppressed and can be safely neglected, leaving six complex amplitudes.

The labels L and R refer to the chirality of the dimuon system. In the limit that the decay is dominated by a vector current $A_{\parallel,\perp,0}^L = A_{\parallel,\perp,0}^R$, $J_{5,6,7} = 0$ and the angular expression collapses to the expression for $B^0 \to J/\psi\, K^{*0}$:

$$\begin{aligned}\frac{\mathrm{d}^3\Gamma[\overline{B}^0 \to J/\psi\,\overline{K}^{*0}]}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi} = \frac{9}{32\pi}\Big[\; & 2|A_0|^2\cos^2\theta_K\sin^2\theta_\ell \;+ \\ & \frac{1}{2}\left(|A_\parallel|^2+A_\perp|^2\right)\sin^2\theta_K(1+\cos^2\theta_\ell) \;+ \\ & \frac{1}{2}\left(|A_\perp|^2-|A_\parallel|^2\right)\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi \;+ \\ & \sqrt{2}\Re e(A_0A_\parallel^*)\sin 2\theta_K\sin 2\theta_\ell\cos\phi \;+ \\ & \frac{1}{\sqrt{2}}\Im(A_0A_\perp^*)\sin 2\theta_K\sin 2\theta_\ell\sin\phi \;+ \\ & \Im(A_\parallel A_\perp^*)\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big]\end{aligned} \tag{1.54}$$

In the limit of $q^2 \gg 4m(\mu)^2$ the factor $\beta_\mu^2 \to 1$ and :

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} = |A_\perp^L|^2+|A_\parallel^L|^2+|A_0^L|^2+|A_\perp^R|^2+|A_\parallel^R|^2+|A_0^R|^2$$

In this limit, $J_2^c = -J_1^c$ and $J_2^s = J_1^s/3$. While the differential decay rate in Eq. (1.52) is defined for the decay of the $\overline{B}^0$ meson, the decay of the B^0 is given in complete analogy by

$$\frac{\mathrm{d}^4\Gamma[B^0 \to K^{*0}\mu^+\mu^-]}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2} = \frac{9}{32\pi}\sum_i \bar{J}_i(q^2)f_i(\cos\theta_\ell,\cos\theta_K,\phi). \tag{1.55}$$

The identical form of this equation compared to Eq. (1.52) is a consequence of our angular definition described in Sec. 1.3.1. Following Ref. [4], it is customary to define CP-averaged observables S_i and CP-violating observables A_i according to

$$S_i = \frac{J_i+\bar{J}_i}{\left(\mathrm{d}\Gamma+\mathrm{d}\bar{\Gamma}\right)/\mathrm{d}q^2} \tag{1.56}$$

$$A_i = \frac{J_i-\bar{J}_i}{\left(\mathrm{d}\Gamma+\mathrm{d}\bar{\Gamma}\right)/\mathrm{d}q^2}. \tag{1.57}$$

The normalisation condition implies $\frac{3}{4}(2S_1^s+S_1^c)-\frac{1}{4}(2S_2^s+S_2^c)=1$. In the limit of massless leptons, the CP-averaged observables are related by $S_2^c=-S_1^c$ and $S_2^s=S_1^s/3$ as discussed above.

Often, the forward-backward asymmetry A_{FB}, and the longitudinal (transverse) polarisation fraction F_L (F_T) are referred to in the literature. These quantities are related to the CP-averaged observables S_i according to:

$$\begin{aligned}A_{\mathrm{FB}} &= \frac{3}{4}S_6^s \\ F_{\mathrm{L}} &= S_1^c = -S_2^c \\ F_{\mathrm{T}} &= 4S_2^s.\end{aligned}$$

In the CP-averaged basis, the $B^0 \to K^{*0}\mu^+\mu^-$ angular distribution is given by,

$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^3(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi} = \frac{9}{32\pi}\Big[\frac{3}{4}(1-F_L)\sin^2\theta_K + F_L\cos^2\theta_K + \frac{1}{4}(1-F_L)\sin^2\theta_K\cos 2\theta_\ell - F_L\cos^2\theta_K\cos 2\theta_\ell + S_3\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi + S_4\sin 2\theta_K\sin 2\theta_\ell\cos\phi + S_5\sin 2\theta_K\sin\theta_\ell\cos\phi + \frac{4}{3}A_{\mathrm{FB}}\sin^2\theta_K\cos\theta_\ell + S_7\sin 2\theta_K\sin\theta_\ell\sin\phi + S_8\sin 2\theta_K\sin 2\theta_\ell\sin\phi + S_9\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big]. \tag{1.58}$$

1.3.3 Interference with other $K^+\pi^-$ states

Eq. 1.58 is valid if the $K^+\pi^-$ system is in a P-wave configuration, as is the case for the $K^{*0}(892)$ vector meson. If the $K^+\pi^-$ system is in a S-wave configuration or in a configuration with higher angular momentum up to $J_{\max}$, the substitutions:

$$A(J=1)^{L,R}_0 \cdot Y^0_1(\theta_K,0) \to \sum_{i=0}^{J_{\max}} A^{L,R}_0(i)\cdot Y^0_i(\theta_K,0) \text{ and} \tag{1.59}$$

$$A(J=1)^{L,R}_{\parallel,\perp} \cdot Y^0_1(\theta_K,0) \to \sum_{i=1}^{J_{\max}} A^{L,R}_0(i)\cdot Y^{-1}_i(\theta_K,0) \tag{1.60}$$

need to be made, where the $Y^m_l(\theta_K)$ are spherical harmonics. The relevant spherical harmonics for S, P and D-wave are

$$\begin{aligned} Y^0_0(\theta_K) &= \frac{1}{2\sqrt{\pi}} \\ Y^0_1(\theta_K) &= \frac{1}{2}\sqrt{\frac{3}{\pi}}\cos\theta_K \\ Y^0_2(\theta_K) &= \frac{1}{4}\sqrt{\frac{5}{\pi}}(3\cos^2\theta_K - 1) \\ Y^{-1}_1(\theta_K) &= \frac{1}{2}\sqrt{\frac{3}{2\pi}}\sin\theta_K \\ Y^{-1}_2(\theta_K) &= \frac{1}{2}\sqrt{\frac{15}{2\pi}}\sin\theta_K\cos\theta_K. \end{aligned}$$

1.3.3.1 S-wave interference

For the decay $B^0 \to J/\psi K^{*0}$ the S-wave fraction was determined to be $(6.4\pm0.3\pm1.0)\%$ in a mass window of $\pm 70\,\mathrm{MeV}$ around the known K^{*0} mass using $1\,\mathrm{fb}^{-1}$ of LHCb data [11]. In the previous publications [8, 9] the presence of an S-wave contribution in $B^0 \to K^{*0}\mu^+\mu^-$ was accounted for by assigning a systematic uncertainty. In this analysis, the S-wave parameters are included in the maximum likelihood fit and treated as nuisance parameters. Therefore,

Eq. (1.52) needs to be modified according to

$$\begin{aligned}\left.\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2}\right|_{S+P} \to \left.\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi\,\mathrm{d}q^2}\right|_{P} + \\ \frac{9}{32\pi}\Big[\, J'^c_1 + J''^c_1\cos\theta_K \;+ J'^c_2\cos 2\theta_\ell \;+ J''^c_2\cos\theta_K\cos 2\theta_\ell \;+ \\ J'_4\sin 2\theta_\ell\sin\theta_K\cos\phi \;+ J'_5\sin\theta_\ell\sin\theta_K\cos\phi \;+ \\ J'_7\sin\theta_\ell\sin\theta_K\sin\phi \;+ J'_8\sin 2\theta_\ell\sin\theta_K\sin\phi\Big]\end{aligned} \tag{1.61}$$

with

$$\begin{aligned} J'^c_1 &= \frac{1}{3}|A^L_{J=0}|^2 + \frac{1}{3}|A^R_{J=0}|^2 \\ J''^c_1 &= \frac{2}{\sqrt{3}}\Big[\Re e(A^L_{J=0}A^{L*}_0) + \Re e(A^R_{J=0}A^{R*}_0)\Big] \\ J'^c_2 &= -\Big[\frac{1}{3}|A^L_{J=0}|^2 + \frac{1}{3}|A^R_{J=0}|^2\Big] \\ J''^c_2 &= -\frac{2}{\sqrt{3}}\Big[\Re e(A^L_{J=0}A^{L*}_0) + \Re e(A^R_{J=0}A^{R*}_0)\Big] \\ J'_4 &= \sqrt{\frac{2}{3}}\Big[\Re e(A^L_{J=0}A^{L*}_\parallel) + \Re e(A^R_{J=0}A^{R*}_\parallel)\Big] \\ J'_5 &= 2\sqrt{\frac{2}{3}}\Big[\Re e(A^L_{J=0}A^{L*}_\perp) - Ree(A^R_{J=0}A^{R*}_\perp)\Big] \\ J'_7 &= 2\sqrt{\frac{2}{3}}\Big[\Im m(A^L_{J=0}A^{L*}_\parallel) - \Im m(A^R_{J=0}A^{R*}_\parallel)\Big] \\ J'_8 &= \sqrt{\frac{2}{3}}\Big[\Im m(A^L_{J=0}A^{L*}_\perp) + \Im m(A^R_{J=0}A^{R*}_\perp)\Big] \end{aligned} \tag{1.62}$$

and

$$\begin{aligned}\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} = \frac{\mathrm{d}\Gamma_P}{\mathrm{d}q^2} + \frac{\mathrm{d}\Gamma_S}{\mathrm{d}q^2} = \Big[|A^L_{J=0}|^2 + |A^L_{J=1,0}|^2 + |A^L_{J=1,\parallel}|^2 + |A^L_{J=1,\perp}|^2 \;+ \\ |A^R_{J=0}|^2 + |A^R_{J=1,0}|^2 + |A^R_{J=1,\parallel}|^2 + |A^R_{J=1,\perp}|^2\Big]\end{aligned} \tag{1.63}$$

The fraction of longitudinal polarisation is given by

$$F_{\mathrm{S}} = |A^L_{J=0}|^2 \cdot \frac{1}{\mathrm{d}\Gamma/\mathrm{d}q^2} \tag{1.64}$$

In the CP-averaged basis, the $B^0 \to K^{*0}\mu^+\mu^-$ angular distribution needs to be changed to include both the S-wave and interference between S- and P-wave, resulting in

$$\begin{aligned}\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{S+P}} &= (1-F_S)\,\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{P}} \\ &+\tfrac{3}{16\pi}\big[F_S\sin^2\theta_\ell + S_{S1}\sin^2\theta_\ell\cos\theta_K \\ &+ S_{S2}\sin 2\theta_\ell\sin\theta_K\cos\phi \\ &+ S_{S3}\sin\theta_\ell\sin\theta)K\cos\phi \\ &+ S_{S4}\sin\theta_\ell\sin\theta_K\sin\phi \\ &+ S_{S5}\sin 2\theta_\ell\sin\theta_K\sin\phi\big].\end{aligned} \tag{1.65}$$

1.3.4 Less form-factor dependent observables

The angular observables can be reparametrized such that leading form-factor uncertainties cancel to first order. The authors of Ref. [12] propose the basis consisting of F_{L} (or A_{FB}) and the observables $P_i^{(\prime)}$ that can be calculated from the observables S_i according to:

$$\begin{aligned}P_1 &= 2\frac{S_3}{1-F_{\mathrm{L}}} \\ P_2 &= \frac{1}{2}\frac{S_6^s}{1-F_{\mathrm{L}}} = \frac{2}{3}\frac{A_{\mathrm{FB}}}{1-F_{\mathrm{L}}} \\ P_3 &= -\frac{S_9}{1-F_{\mathrm{L}}} \\ P_4' &= \frac{S_4}{\sqrt{F_{\mathrm{L}}(1-F_{\mathrm{L}})}} \\ P_5' &= \frac{S_5}{\sqrt{F_{\mathrm{L}}(1-F_{\mathrm{L}})}} \\ P_6' &= \frac{S_7}{\sqrt{F_{\mathrm{L}}(1-F_{\mathrm{L}})}} \\ P_8' &= \frac{S_8}{\sqrt{F_{\mathrm{L}}(1-F_{\mathrm{L}})}}.\end{aligned}$$

1.4 R_{K^*} measurement

In the SM, the gauge bosons couple equally to the different flavours of lepton. The Higgs boson instead couples differently and is responsible for the mass spectrum of the leptons. Hiller and Krüger [13] proposed to test this lepton universality measuring the ratio R_H of the decay rates of B mesons into final states with leptons.

$$R_H = \frac{\Gamma[B \to H\mu^+\mu^-]}{\Gamma[B \to He^+e^-]} \tag{1.66}$$

where H can be an inclusive state containing a s quark (X_s) or a s quark resonance: ϕ, K, K^*.

These ratios are expected to be $R_H = 1$, in the case of massless leptons. Taking into account the lepton mass, giving an effects of order $\frac{\mathrm{m}_\mu^2}{m_b^2}$, the prediction [13, 14]of the SM in the full q^2 range are:

$$R_K = 1.0003 \pm 0.0001,$$
$$R_{K^*} = 0.991 \pm 0.002$$

Given the high precision of the predictions, the measurement of the R_H allows to search in a model independent way the existence of non Standard Model scalar and pseudo scalar interactions with couplings depending on the lepton flavour.

After the previous measurement from BaBar [15] and Belle [16], in 2014 the LHCb experiment, has measured R_K [17] in the $1 < q^2 < 6$ range:

$$R_K[1,6] = 0.745^{+0.090}_{-0.074}(\text{stat}) \pm 0.036(\text{syst}),$$

which differs from the SM prediction by 2.6σ, as shown in Figure 1.4.

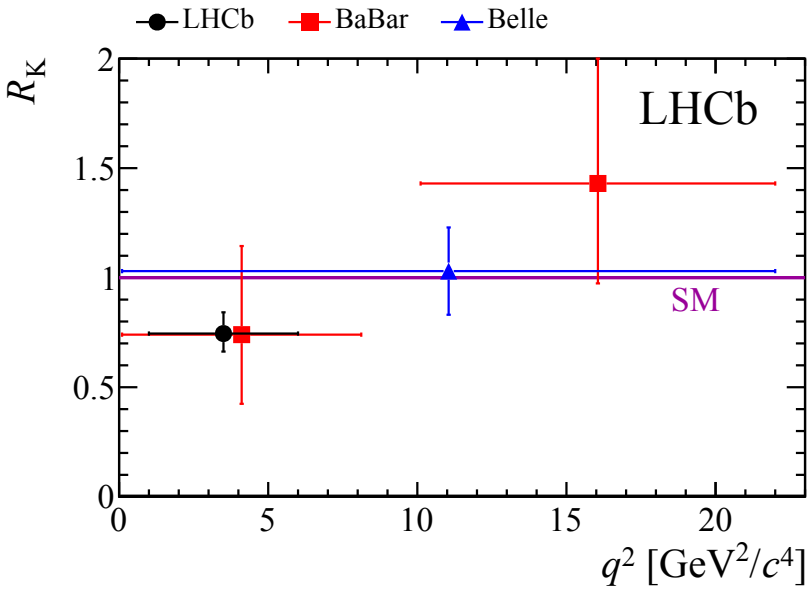


Figure 1.4: R_K ratio measurements by the LHCb (black dot), BaBar (red square) and Belle (blue triangle) experiments.

This result is not an evidence of presence of new physics by itself, but can be related with the results from BaBar on $B \to D^{(*)}\tau\nu$ decays [18, 19] which have also shown hints of a violation of lepton universality. A recent measurement of $R(D^*)$ by the LHCb experiment [20] has confirmed the deviation of 2.1σ from the SM seen by BaBar. These results have led to many theoretical speculations. In this context, a measurement of the R_{K^*} ratio is expected to confirm or not these hints of lepton non-universality and will help to distinguish between alternative New Physics models [21]. At the moment the analysis is being performed at LHCb with 3 fb^{-1} of data, and the current status is presented in Chapter 6. The meauserement is done in three q^2 regions. This choice is related to the different diagrams contributing in each region. At low q^2, the $b \to s\ell^+\ell^-$ decays are dominated by the photon pole and the C_7 Wilson coefficient. In the central bin, $1 < q^2 < 6$ GeV$^2/c^4$, the decay is dominated by an interference between C_7 and C_9. On the high-q^2 region, we have a mixture of C_9 and C_{10}.

The measurement of R_{K^*} in each q^2-region will be obtained by the following formula:

$$R_{K^*}(q^2_{\text{min}}, q^2_{\text{max}}) = \frac{\int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \frac{d\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)}{dq^2}}{\int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \frac{d\mathcal{B}(B^0 \to K^{*0}e^+e^-)}{dq^2}} dq^2$$

The C_7 Wilson coefficient has been also largely tested through measured of $B \to K^*\gamma$, [22] and is in good agreement with the SM. So we do not expect large effects of new physics to show up in the first q^2 bin, but the measurement in the low q^2 region will anyway constitute a solid validation of the analysis technique for the other regions.

REFERENCES

[1] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531.

[2] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[3] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[4] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of $B \to K^* \mu^+ \mu^-$ Decays in the Standard Model and Beyond*, JHEP **01** (2009) 019, arXiv:0811.1214.

[5] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetries in the decays $B^0 \to K^{*0}\mu^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$*, JHEP **2014** (2014), no. 9 , arXiv:1408.0978.

[6] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetry of $B \to K^{(*)}\mu^+\mu^+$ decays*, JHEP **06** (2014) 133, arXiv:1403.8044.

[7] C. Bobeth, G. Hiller, and G. Piranishvili, *CP Asymmetries in bar $B \to \bar{K}^*(\to \bar{K}\pi)\bar{\ell}\ell$ and Untagged $\bar{B}_s$, $B_s \to \phi(\to K^+K^-)\bar{\ell}\ell$ Decays at NLO*, JHEP **0807** (2008) 106, arXiv:0805.2525.

[8] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$*, JHEP **08** (2013) 131, arXiv:1304.6325.

[9] LHCb collaboration, R. Aaij *et al.*, *Measurement of form-factor-independent observables in the decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Rev. Lett. **111** (2013) 191801, arXiv:1308.1707.

[10] T. Blake and C. Langenbruch, *Angular conventions for the decays $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$*, LHCb-INT-2012-021, 2012.

[11] LHCb collaboration, R. Aaij *et al.*, *Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays*, Phys. Rev. **D88** (2013) 052002, arXiv:1307.2782.

[12] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, *Optimizing the basis of $B \to K^*\ell^+\ell^-$ observables in the full kinematic range*, JHEP **1305** (2013) 137, arXiv:1303.5794.

[13] G. Hiller and F. Kruger, *More model independent analysis of $b \to s$ processes*, Phys. Rev. **D69** (2004) 074020, arXiv:hep-ph/0310219.

[14] C. Bobeth, G. Hiller, and G. Piranishvili, *Angular distributions of anti-B —> K anti-l l decays*, JHEP **12** (2007) 040, arXiv:0709.4174.

[15] BaBar, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \to K^{(*)}l^+l^-$*, Phys. Rev. **D86** (2012) 032012, arXiv:1204.3933.

[16] J. Seely *et al.*, *New Measurements of the European Muon Collaboration Effect in Very Light Nuclei*, Phys. Rev. Lett. **103** (2009) 202301.

[17] LHCb, R. Aaij *et al.*, *Test of lepton universality using* $B^+ \to K^+\ell^+\ell^-$ *decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482.

[18] BaBar, J. P. Lees *et al.*, *Evidence for an excess of* $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ *decays*, Phys. Rev. Lett. **109** (2012) 101802, arXiv:1205.5442.

[19] BaBar, J. P. Lees *et al.*, *Measurement of an Excess of* $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ *Decays and Implications for Charged Higgs Bosons*, Phys. Rev. **D88** (2013), no. 7 072012, arXiv:1303.0571.

[20] LHCb, R. Aaij *et al.*, *Measurement of the ratio of branching fractions* $\mathcal{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau))/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$,arXiv:1506.0861.

[21] G. Hiller and M. Schmaltz, *Diagnosing lepton-nonuniversality in* $b \to s\ell\ell$, JHEP **02** (2015) 055, arXiv:1411.4773.

[22] BABAR Collaboration, B. Aubert *et al.*, *Measurement of* $B \to K^*(892)\gamma$ *Branching Fractions and* CP *and Isospin Asymmetries*, Phys. Rev. Lett. **103** (2009) 211802.

Chapter 2

The LHCb experiment

Sommaire

The LHCb experiment at the Large Hadron Collider LHC of the CERN laboratory is a international collaboration of more than 1121 physicists from 68 institutions and 16 different countries. In France five laboratories of the IN2P3 institute are involved in the project. The experiment participates in the challenge of finding answers to the fundamental particle physics open questions, with a particular focus on the problem of the matter-antimatter asymmetry and on the indirect searches for new physics effects in b and c-hadrons decays. After recalling the main features of the LHC accelerator and the summary of its operation during the first period of the data taking, this chapter will present the experimental apparatus used to collect the data on which the analyses presented in this thesis have been performed.

2.1 The LHC

The *Large Hadron Collider* (LHC) is currently the biggest and most powerful particle accelerator in the world. This circular accelerator is based at CERN (European Organisation for Nuclear Research), in the French-Switzerland border, near Geneva. The LHC uses the 27

kilometres underground ring that was holding the *Large Electron Positron collider* (LEP) in the 90's. It has been conceived to accelerate protons up to an energy reaching 6.5 TeV per nucleon, in two beams of opposite directions which collide in four different points inside the tunnel, where the main experiments are localized:

- two general purpose detectors, ATLAS [1] (*A Thoroidal Lhc ApparatuS*) and CMS [2] (*Compact Muon Solenoid*), have been conceived to search for the Higgs boson, accomplish some precision measurements of the standard model and also look for phenomenon beyond the standard model.
- the detector LHCb [3](*LHC beauty*), optimized to study the flavour physics.
- the detector ALICE [4] (*A Large Ion Collider Experiment*), dedicated to the study of the quark-gluon plasma in the heavy ions collisions.

Three additional experimental facilities, smaller in size, are installed. TOTEM is an experiment dedicated to the measurement of total cross section, elastic scattering and diffractive processes at LHC energies. LHCf (LHC-forward) is dedicated to measure the production cross section of neutral particles in the forward direction in pp and heavy ions collisions, to better constrain models of energetic showers. MoEDAL (Monopole and Exotics Detector At the LHC) is dedicated to the search for magnetic monopole and other highly ionizing stable massive particles, and is installed inside the LHCb pit. A schematic overview of the experiments, the LHC and the CERN accelerator complex is given in the figure 2.1.

2.1.1 Summary of the first period of work (Run I)

The first data taking period has started in 2010 and has lasted until the beginning of 2013. LHCb has recorded a luminosity of 1.1 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ (2010-2011), and 2.08 fb^{-1} at $\sqrt{s} = 8\,\text{TeV}$ (2012). During the Run I, LHCb measured with the 2011 dataset the cross-section at $\sqrt{7}$ TeV to be $284 \pm 20 \pm 49$ μb [5]. A total of 26×10^{10} of $b\bar{b}$ and 59×10^{11} $c\bar{c}$ pairs were produced within the LHCb acceptance.

2.2 The LHCb detector

The LHCb [3] detector is a single-arm forward spectrometer designed to perform precision measurements of b and c hadrons decays. LHCb has an angular coverage in the [10 mrad, 300 mrad (250 mrad)] range in the bending (non-bending) plane, corresponding to a pseudorapidity[1] region of $2 < \eta < 5$. This geometry is adopted since the $b\bar{b}$ pairs are mostly produced in the forward (and backward) direction at the LHC, as shown by PYTHIA simulation in Fig.2.2.

The layout of the LHCb detector can be seen in Fig. 2.3. The right-handed coordinate system adopted by LHCb is defined has the z axis along the beam axis and the y axis along the vertical.

The LHCb detector includes several sub-detectors providing measurement of the trajectory of the particles (vertex locator, trackers) and their identifications (RICH, calorimeters and muon chambers). The whole detector is describe in this section as well as the trigger system.

[1] The pseudorapidity is defined as $\eta = -\ln[\tan\frac{\theta}{2}]$, where θ is the angle between particle momentum and the beam axis.

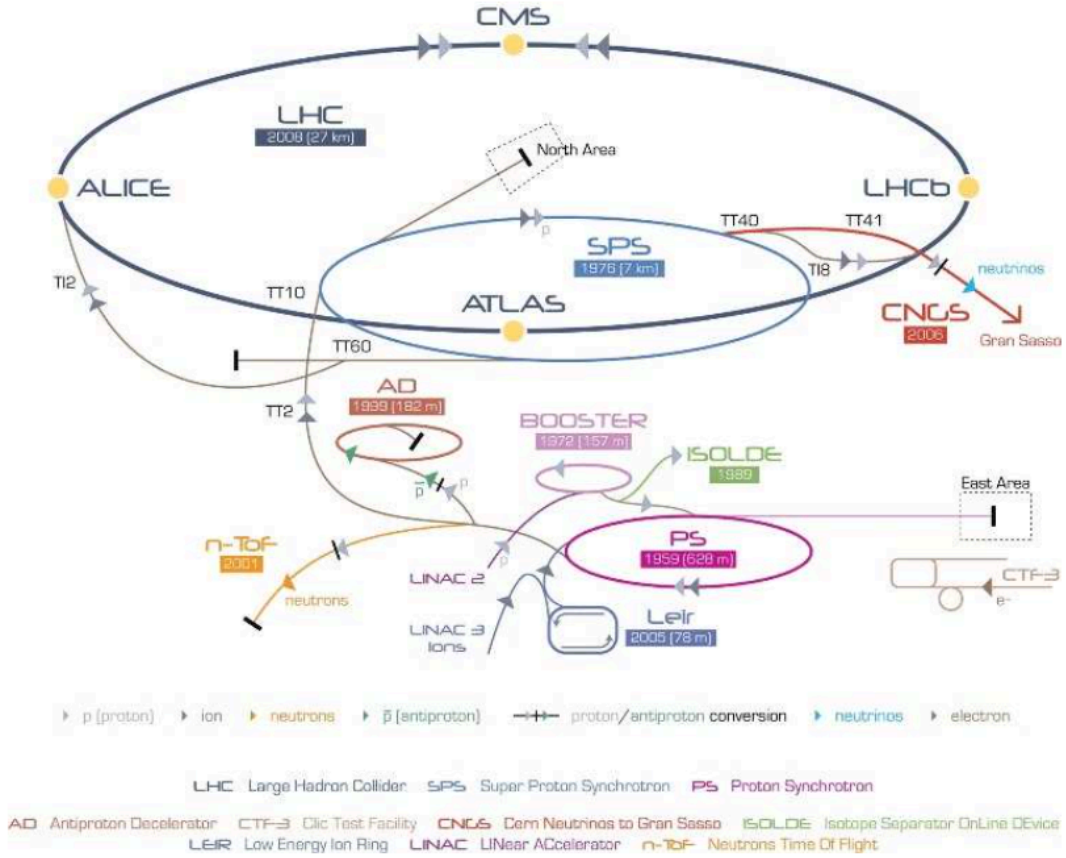


Figure 2.1: The CERN accelerator complex.

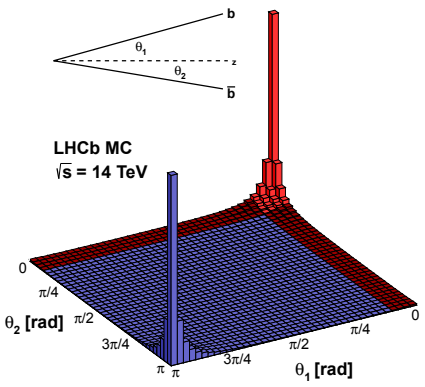


Figure 2.2: Polar angles correlation of the b and the $\overline{b}$, obtained from a PYTHIA 8 simulation at $\sqrt{14}$ TeV. The LHCb acceptance is represented by the red area.

s

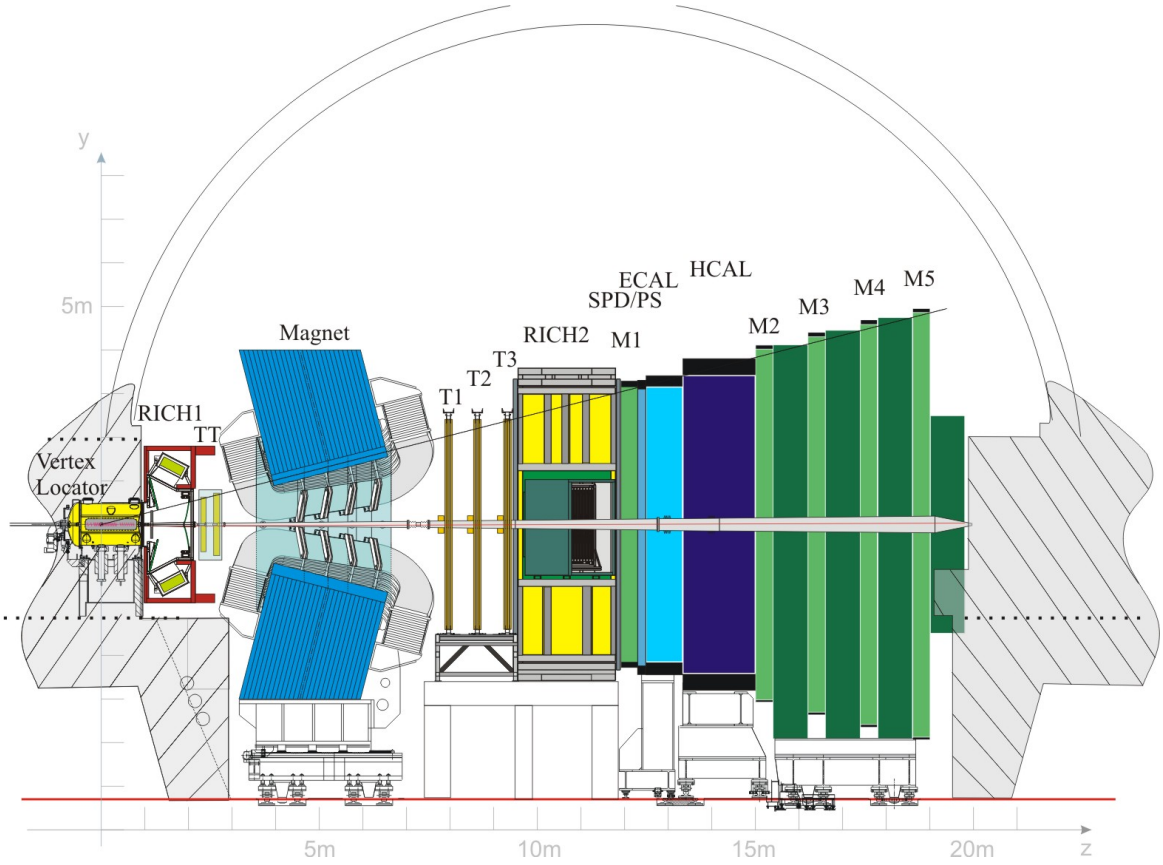


Figure 2.3: Layout of the LHCb detector, with the different sub-detectors and the right-handed coordinate system.

2.2.1 The tracking system

In LHCb the tracking system is composed of five subsystem: the VErtex Locator (VELO) placed around the region where the beams collide, a magnet, the Trigger Tracker (TT) in front of the magnet and three tracking stations (T1-T3) behind the magnet. The VELO and TT are silicon microstrip detectors, while T1-T3 use silicon microstrip in their inner parts and straw tubes in their outer parts.

2.2.1.1 The VErtex LOcator

The VELO is the sub-detector which is closest to the interaction point. It is conceived to provide a precise primary and secondary vertex determination. This determination allows to measure the decay lifetimes τ, of the b and c hadrons decays (approximately 10^{-12} s for B mesons and $0.5-1\times10^{-12}$ s for D mesons). The VELO is composed of 21 stations of silicon modules (Figure. 2.4) placed along the beam direction. Each station has 2 types of modules : R-sensors to measure the radial distance to the beam axis, and ϕ-sensors to measure the azimuthal coordinates of the charged particles. Two additional $pile-up$ stations are placed upstream of the interaction point to allow a fast determination of the number of primary vertices as well as the backward charged tracks multiplicity. Each sensor has an external radius of 42 mm and an internal one of 8 mm, and they are made up of 2048 silicon strips. The two halves of the VELO are retractable, allowing to have a larger opening, up to 6 cm, to protect the sensor during the beams injection. During normal condition, they overlap for alignment purposes, and cover an acceptance of $1.6<\eta<4.9$.

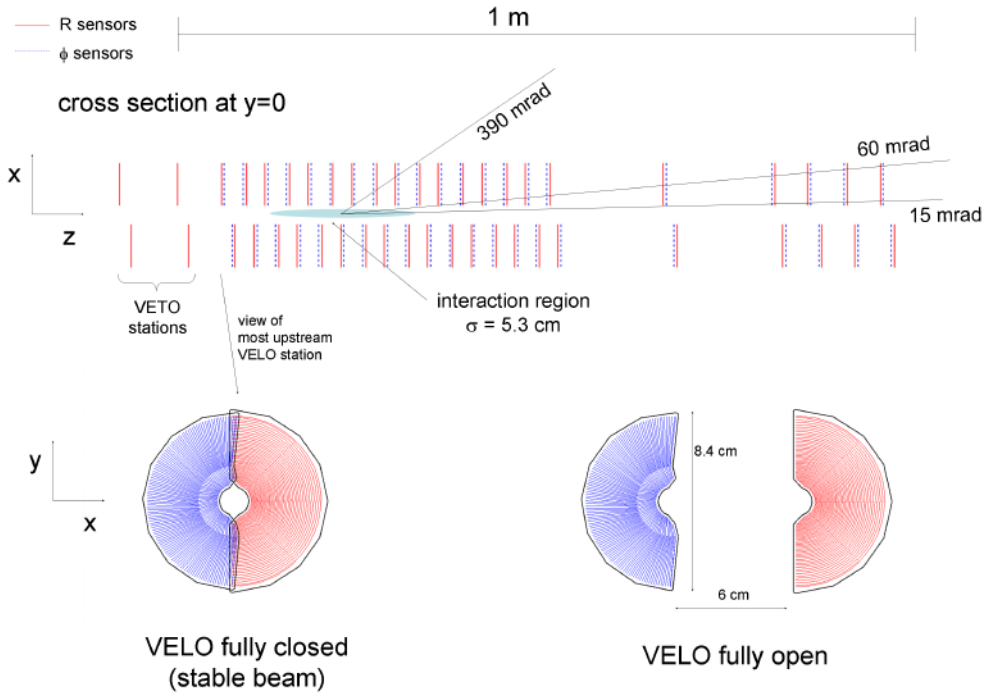


Figure 2.4: Cross section in the (x,y) plane of the VELO silicon sensors, at $y=0$, with the detector in the fully closed position. The two drawings on the bottom illustrate the closed and open positions. The two pile-up veto stations are located upstream of the VELO sensors.

2.2.1.2 The tracking system and magnet

The magnet

The dipole magnet of LHCb, shown in Figure 2.5, is a non superconducting dipole with an integrated field of 4 Tm, required for high momentum resolution. It's composed of two identical coils of conical saddle shape, placed mirror-symmetrically to each other in the magnet yoke.

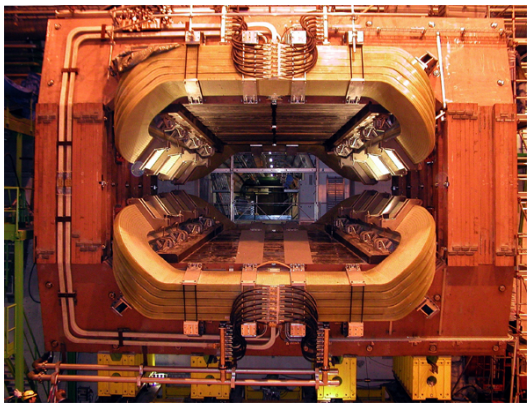


Figure 2.5: Picture of the LHCb dipole magnet installed in the cavern.

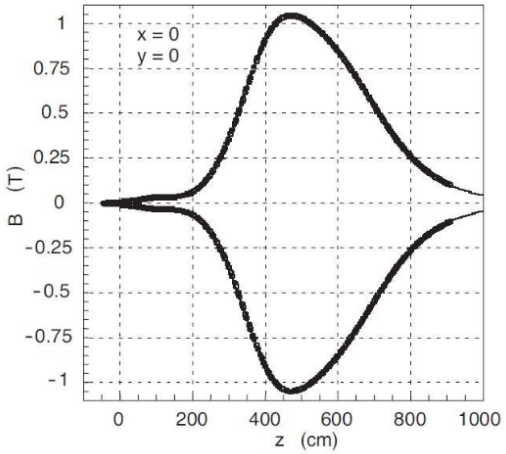


Figure 2.6: Magnetic field B_y along the z axis for both magnet polarities.

A magnetic field mapping measurement has been performed for the three spatial components. In Figure 2.6 it is shown for the y component of the field as an example, for both polarities. The non-uniformities of the field are at the order of 1%. The polarity of the magnet can be reversed in order to study systematics effects due to possible left-right detection asymmetries.

The silicon tracker: Tracker Turicensis and Inner Tracker

The silicon tracker is the generic name for two detectors: the Tracker Turicensis (TT) located upstream of the magnet and the Inner Tracker (IT) located downstream of the magnet. Scheme of an IT and a TT detector layer are shown in Figure 2.7 and 2.8 respectively. They both use silicon microstrip sensors with a strip pitch of about 200 µm providing a 50 µm single hit spatial resolution. Each station has four layers disposed along the $x-u-v-x$ axis, where x refer to the vertical and the u and v are tilted around the z axis by $-5°$ and $+5°$ respectively.

The TT is composed of four tracking station layers located upstream of the magnet and downstream of the RICH1 detector. It is a rectangular area of 150 cm wide and 130 cm high, covering the full LHCb acceptance.

The IT systems is located downstream of the magnet covering the inner region of the three tracking stations T1, T2 and T3. They are 120 cm × 40 cm cross-shaped planes.

Outer Tracker

The Outer Tracker (OT) is a drift time system for precise momentum measurement and track detection, Figure 2.9. The three stations T1, T2 and T3 are placed downstream the magnet in the same plane as the IT. Each station is made of four layer of 64 gas-tight straw-tubes of 4.9 mm of diameter. The four layer are also disposed along the $x-u-v-x$ geometry used by

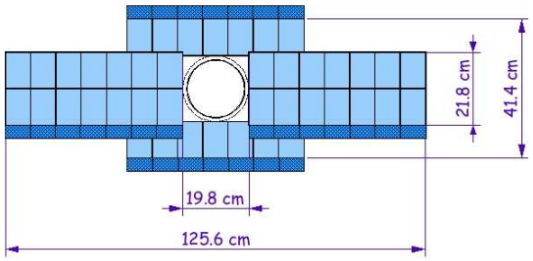


Figure 2.7: Scheme of an IT station.

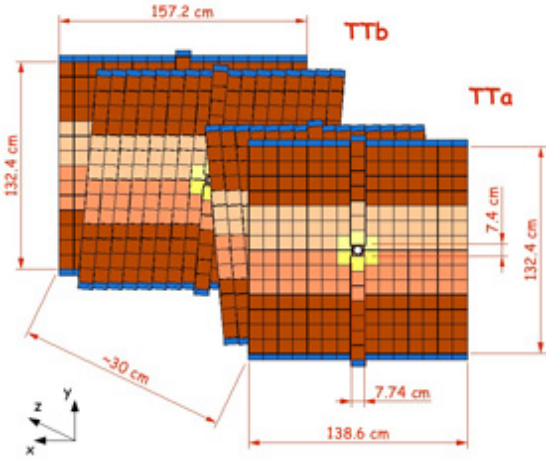


Figure 2.8: Layout of a TT detector module

the other tracking stations. The gas in the straw-tubes is a mixture of 70% Argon and 30% CO_2 which provides a fast drift time, below 50 ns, and a good drift-coordinate resolution of 200 µm.

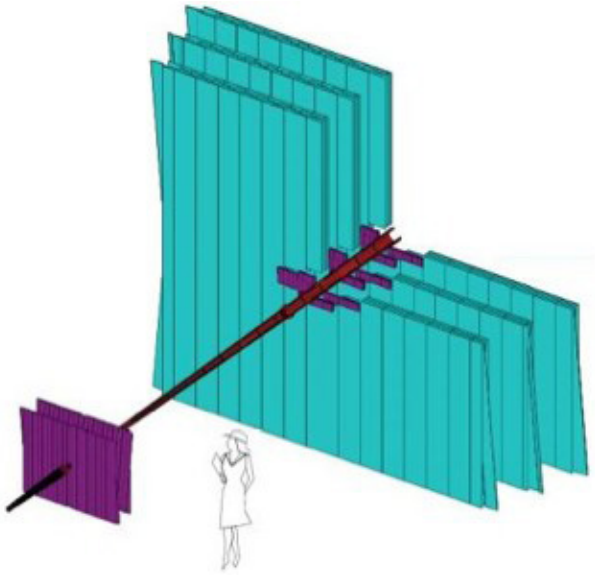


Figure 2.9: Arrangement of OT straw-tube modules in layers and stations (cyan) around the IT stations (purple).

2.2.2 The particle identification system

2.2.2.1 The RICH detector

The precise particle identification, mainly needed to discriminate between the different hadronic channels, is performed by two dedicated Ring Imaging Cherenkov (RICH) systems. For tracks with low momentum range, $1-60\,\mathrm{GeV}/c$, the particle identification is provided by the first RICH detector (RICH1), Figure 2.10, which is located upstream the magnet. Instead, for higher momenta particles from ≈ 15 to $100\,\mathrm{GeV}/c$, is the second RICH (RICH2), Figure 2.11,

located downstream the magnet, that gives the particle identification information. The RICH1 is made of aerogel and fluorobutane C_4F_{10} while the RICH2 uses CF_4 as radiator material. The choice of the radiators is illustrated in Figure 2.12. The RICH2 has a limited acceptance in comparison with the LHCb detector, 15 mrad to 120mrad (100 mrad) in the bending (non-bending) plane, but covers the area where the high momentum particles are located.

A charged particle passing through the radiator with a velocity $c\beta$ greater than the speed of light c/n in the material, emits Cherenkov photons in shape of a cone around the direction of propagation. The angle of the cone, called Cherenkov angle, is expressed as $\cos\theta_{\text{Cherenkov}} = \frac{1}{n\beta}$ and is related to the mass and momentum of the particle. In both RICH1 and RICH2, the Cherenkov light emitted by the particle is reflected by a combination of mirrors and directed to the Hybrid Photon Detectors (HPDs) installed outside the LHCb acceptance, where it is collected.

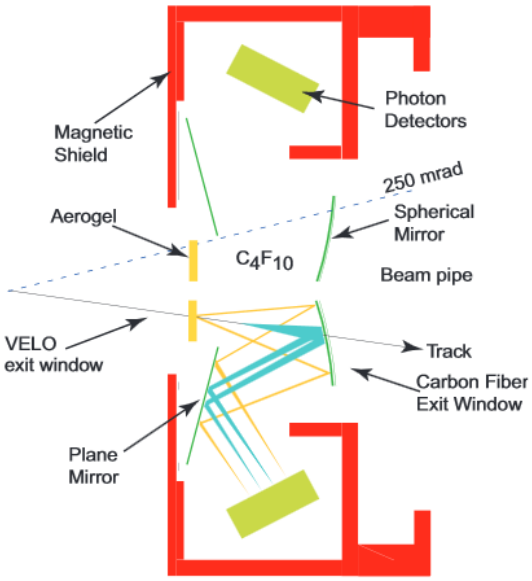


Figure 2.10: Side view schematic layout of the RICH1 detector

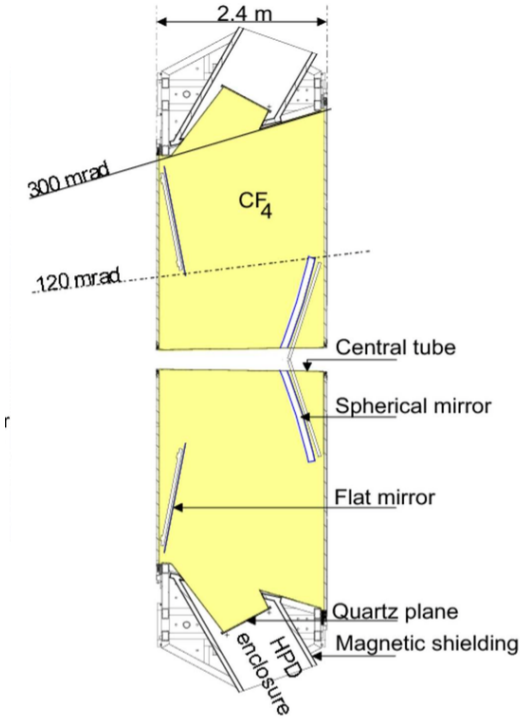


Figure 2.11: Side view schematic layout of the RICH2 detector

2.2.2.2 The calorimeters

The LHCb calorimeters are conceived to identify the electrons, photons and hadrons and to measure their energies and direction. The calorimeters are also used in the hardware trigger to select particles with high transverse energy. Four subsystems compose the LHCb calorimeter: a Scintillator Pad Detector (SPD), a PreShower detector (PS) and finally an Electromagnetic (ECAL) and Hadronic (HCAL) calorimeters, shown in Figure 2.13. All the four detectors work with the shashlik technology: they are sampling devices using scintillator material, and the scintillating light is transmitted to photomultiplier tubes (PMT) by wavelength-shifting (WLS)

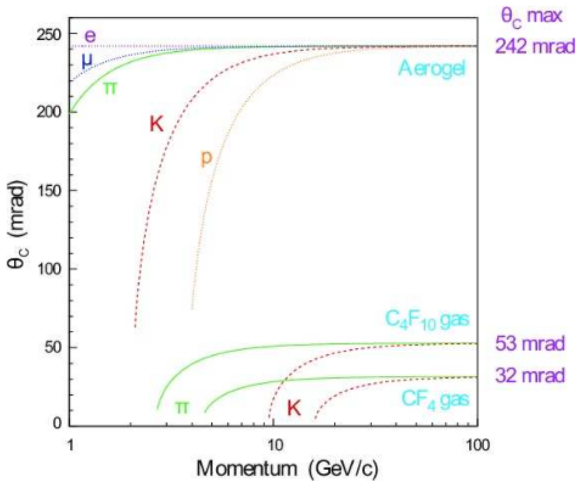


Figure 2.12: Cherenkov angle as function of the particle momentum for different RICH radiators and different particle masses.

fibres. The electronics of this system of sub-detectors is design to be as fast as possible for trigger purposes.

The SPD and PS

The SPD and PS are the first component of the calorimeters. They are composed of a single 15 mm thick scintillators pads with a Pb layer equivalent to $2.5X_0$ placed in between them. To match the ECAL segmentation, the SPD and PS are divided into three regions: Inner (composed of 3072 cells), Middle (3584) and Outer (3376). The WLS fibres are read out by multianode photomultiplier tube (MAPMT), which makes them fast multi-channel detectors.

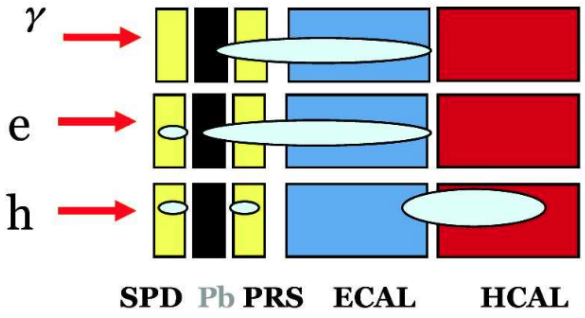


Figure 2.13: Shower development in the calorimeter system for different particle types.

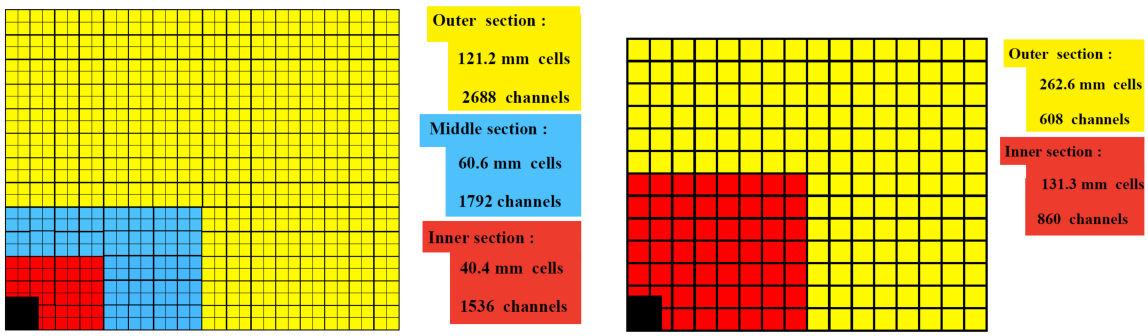
The ECAL

The ECAL is placed at 12.5 m of the interaction point, it is 7.8 m wide and 6.3 m high. Is made of 66 successive layers of 2 mm thick lead and 4 mm thick scintillator. It is 84 cm deep, equivalent to 25 radiation lengths, which guarantees full containment of high energy photon showers. As the hit density is a function of the distance to the beam pipe, and it varies by two order of magnitude over the calorimeters, this sub-detector is divided into three regions. Close

to the z axis, where the particle multiplicity is highest, the detectors cells are smallest, in the Outer region the cells are largest. This design choice provides a modest energy resolution, but on the other hand ensures fast response, uniformity, acceptable radiation resistance and reliability. The energy resolution of the ECAL is $\frac{\sigma_E}{E} = \frac{(9\pm0.5)\%}{\sqrt{E}} \oplus (0.8 \pm 0.2)\%$, where E is expressed in GeV.

The HCAL

The HCAL is placed at 13.33 m of the interaction point, it is 6.8 m wide and 8.4 m high and 1.65 m depth, equivalent to 5.6 interaction lengths. It employs a non-typical structure where the scintillating tiles are arranged parallel to the beam axis and it is transversally segmented in two regions, with double cell size in the Outer region with respect to the Inner one. The HCAL uses 1 cm thick tiles of iron alternated with 3 mm thick scintillator layers. The resolution is $\frac{\sigma_E}{E} = \frac{(69\pm5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%$, where E is expressed in GeV.



(a) ECAL (b) HCAL

Figure 2.14: Cell sizes for the ECAL (left) and for the HCAL (right)

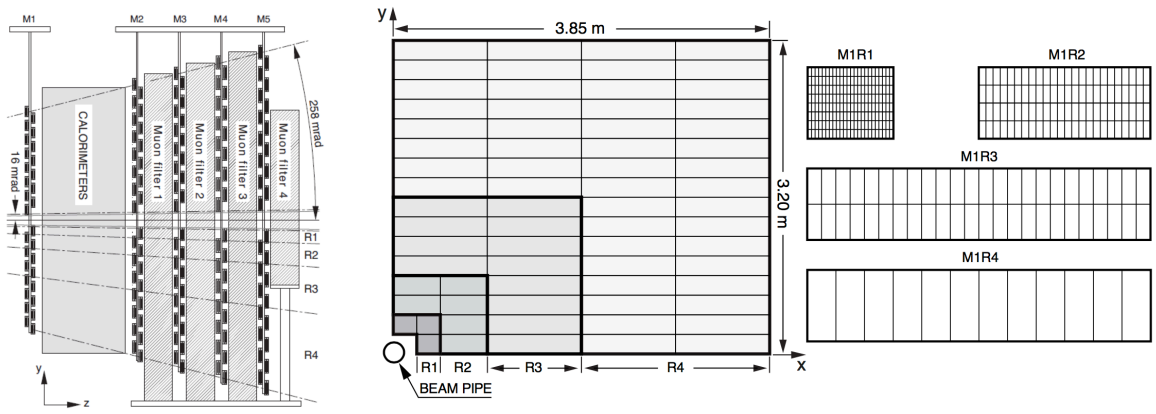
2.2.2.3 The muon chambers

The LHCb muon system provides an essential trigger information as well as offline particle identification. There are five rectangular-shaped muon chamber, placed along the beam axis. The M1 station is located upstream of the SPD, to provide a better transverse momentum measurement at the trigger level, while M2-M5 are downstream of the HCAL. Iron absorbers of 80 cm thickness are placed between the four last stations in order to select penetrating muons.

To have a rather constant particle flux and channel occupancy over the muon detector surface, each station is divided in four regions (R1,R2,R3,R4), with segmentation scale ratios of 1:2:4:8. They are all multi-wire proportional chambers (MWPC), except the R1 regions in M1, which use triple-GEM detectors. Muon stations allow muons to be reconstructed with a 20 % of p_T resolution, sufficient for the L0 trigger. Trajectories from the tracking stations are extrapolated to the muon system and if hits in the muon chambers are found, such tracks are identified as muon candidates. In 2012 the muon identification efficiency was excellent: 97% with only 1-3% pion-to-muon mis-identification probability.

2.2.3 Trigger

The trigger system reduces the rate of the events written on tape from 40 MHz down to few kHz, and it is based on two main levels: the Level-0 trigger (L0), which is hardware based, and



(a) (b)

Figure 2.15: (left) Side view of the muon stations. (right) Logical pad arrangement

the High Level Trigger (HLT), which is software based. The scheme of LHCb trigger system adopted in the 2011-2012 data taking period is presented in Figure 2.16.

The trigger uses the kinematics properties of b hadron decays to identify events in which $b\bar{b}$ pairs are produced. The rate of muon production in pp interactions is considerably smaller than that of hadrons. Muons have also higher efficiency than hadrons. Therefore, the trigger strategy for the $B^0 \to K^{*0}\mu^+\mu^-$ analyses discussed in this thesis is based on the muons.

2.2.3.1 First level

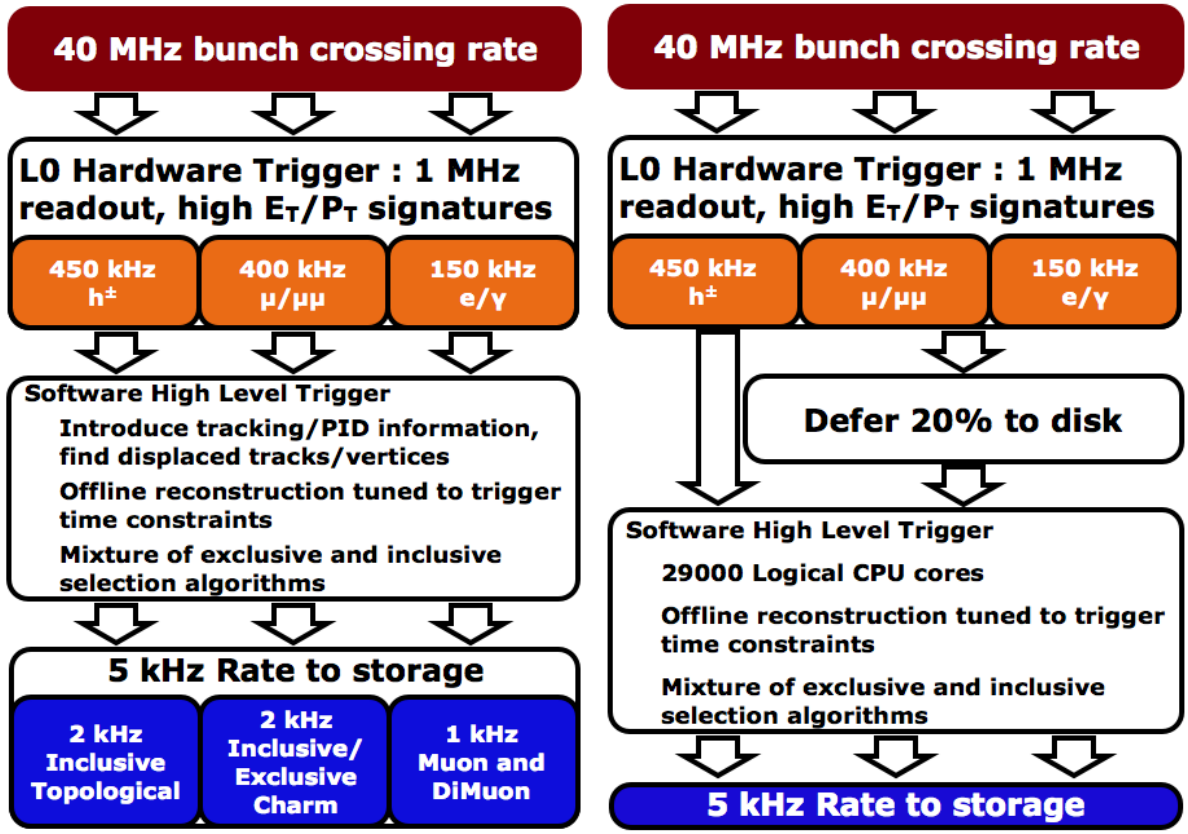
Most of the detector components of LHCb can be read out at a maximum rate of 1 MHz, while the proton bunch crossing frequency is 40 MHz. The Level-0 (L0) is a hardware trigger based on custom made front-end electronics. It operates synchronously with the 40 MHz bunch crossing frequency and reduces the rate to 1 MHz. The L0 exploits the fact that heavy b hadrons have large masses, so their decay products are particles with large transverse momentum (p_{T}) or transverse energy ($\mathrm{E_T}$). The L0 trigger reconstruct and select the highest E_T hadron, electron or photon in the calorimeter systems and the two highest p_T muons in the muon chambers.

The calorimeter trigger system uses all calorimeter stations, forms clusters by adding 2×2 cells and computes the transverse energy of these clusters and select the highest E_T candidates. The transverse energy threshold is $E_T > 3.5$ GeV for hadrons and $E_T > 2.5\,\mathrm{GeV}$ for an electromagnetic shower.

The muon hardware trigger requires one muon with $p_T > 1.48$ GeV/c or two muons with $\sqrt{p_{T,1} \cdot p_{T,2}} > 1.296$ GeV/c and there are no requirement that the two muons have opposite charges.

The thresholds mentioned before refer to the 2011 data taking period. In 2012 the thresholds had to be changed to adapt to the higher collision energy and higher luminosity. They were raised to: 3 GeV for the hadronic trigger, 3 GeV for the electromagnetic trigger, 1.76 GeV/c for the single muon trigger and 1.6 GeV/c for the di-muon trigger.

In addition, a few global quantities are extracted by the L0, as the total number of tracks which is obtained counting the number of hits in the SPD



(a) Trigger system used in 2011

(b) Trigger system with deferring used in 2012

Figure 2.16: Scheme of the different trigger used during 2011 (2.16(a)) and 2012 (2.16(b)) data taking period.

2.2.3.2 HLT level

The software trigger consists of a C++ application reducing the 1MHz output rate of the L0 trigger to 5kHz to be stored on disk. It is subdivided into two main components, HLT1 and HLT2, to optimise the use of the computation power of the computer farm.

HLT1

The first level of the software trigger (HLT1) performs a partial event reconstruction by using the information from the VELO reconstruction algorithm realizing a full 3D pattern recognition. Tracks are selected according to their impact parameter (IP) and their track quality. Then the selected tracks are extrapolated to the tracking stations, where their momentum are determined and a minimum momentum cut is applied. The output rate at this stage is around 30 kHz.

HLT2

After the HLT1 trigger the output rate is low enough to perform a full event reconstruction. VELO tracks are built using a seeding algorithm, and their measured momentum from the tracking stations is required to be $p > 5$ GeV/c with $p_T > 0.5$ GeV/c. An identification algorithm is applied for electrons and muons by matching tracks in ECAL and muon stations. In the HLT2,

the event rate is sufficiently low to run topological triggers, which look for 2-,3- or 4-body objects using the track quality, the IP and the distance of closest approach (DOCA). The HLT2 contains a series of *lines*, each of them selecting a particular types of events. The bandwidth granted to each line is the main constraint in their design, and is modulated according to the LHCb physics program priorities. This flexibility is one of the strong points of this highly performant trigger. In 2012 the so-called "deferred HLT", Figure 2.16(b), was implemented: the idea was to take profit of the time between two LHC fills to process the data in the HLT trigger. This increase the computing resources available over time and so allows to increase the output trigger rate. About 20 % of the data selected by the L0 trigger in a fill were temporary written to disk, waiting to be processed by the HLT during the next inter-fill period.

2.2.3.3 Categories of the trigger decisions: TIS, TOS, Dec

At any level of the trigger, an event can be classified in three non-exclusive categories:

- Trigger On Signal (TOS) : Event which are triggered in the signal decay independently of the presence of the rest of the event. The TOS criterion is satisfied if there exists at least one trigger object all of whose tracks have overlap with the signal.
- Trigger Independent of Signal (TIS) : Events which are triggered independently of the presence of the signal. In order for an event to be TIS, there must exist at least one trigger object which does not have any overlap with the signal. TIS events are trigger unbiased except for correlations between the signal decay and the rest of the event.
- Trigger Decision (Dec): Events which are triggered either by signal trigger (TOS) or by the trigger independent of the signal (TIS) without separating these two categories.

REFERENCES

[1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008), no. 08 S08003.

[2] CMS Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008), no. 08 S08004.

[3] LHCb Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008), no. 08 S08005.

[4] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008), no. 08 S08002.

[5] LHCb collaboration, *Measurement of* $\sigma(pp \to b\bar{b}X)$ *at* $\sqrt{s} = 7$ TeV *in the forward region*, Physics Letters B **694** (2010), no. 3 209 .

Chapter 3

Selection of events for the $B^0 \to K^{*0}\mu^+\mu^-$ angular analysis

Sommaire

The first part of the analysis measuring the angular observables described in Section 1.3.2 is the search for $B^0 \to K^{*0}\mu^+\mu^-$ candidates among all the events collected by the detector. It is important to elaborate an efficient event selection procedure that, while keeping the most of signal events, rejects most of the background events, as they are expected to pollute the angular distribution and to potentially affect the angular measurements. Here we describe this selection procedure, comprising several steps. The first is the choice of trigger lines that allowed to register the interesting events during the data taking, described in Section 3.2. It is followed by a generic loose preselection, named "stripping", presented in Section 3.3, aimed to reduce the amount of data to be analysed with a very high efficiency, and common to a different analyse with two muons in the final state. A set of additional loose requirements constitutes the preselection described in Section 3.4, common to the angular analysis [1] and the S-wave analysis [2]. After these preliminary and more generic steps, the selection focuses on reducing as much as possible the two main background categories affecting the identification of $B^0 \to K^{*0}\mu^+\mu^-$ decays: the peaking background and the combinatorial background. The peaking background comes from exclusive b-hadron decays that are badly reconstructed and follow into the signal region: dedicated particle-identification criteria and invariant mass cuts can reduce their amount, as shown in Section 3.5. Instead, a multivariate technique combining a set of kinematic and identification variables, described in Section 3.6, is used to considerably reduce the level of the combinatorial background, which originates from random combinations of tracks faking the signal. This selection of events, also described extensively in Ref. [3], improves the performances of the selection strategy of the previous analysis [4].

3.1 Data and simulation

The angular analysis described in this thesis is based on data corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$, collected by the LHCb experiment in pp collisions. The dataset comprises $1\,\text{fb}^{-1}$ of data collected at $\sqrt{s} = 7$ TeV in 2011 and $2\,\text{fb}^{-1}$ of data collected at $\sqrt{s} = 8$ TeV in 2012.

Monte Carlo simulated $B^0 \to K^{*0}\mu^+\mu^-$ events, generated according to a phase-space model, are used to determine the efficiency of the selection and reconstruction. The ratio of possible sources of pollution from peaking backgrounds are estimated using several simulated samples of exclusive b-hadron decays. The simulated samples have an approximately equal mix of events generated using Pythia 6 [5] and Pythia 8. The samples used in this analysis are summarised in Tab. 3.1.

3.2 Trigger

Candidates are required to pass the trigger requirements listed in Table 3.2, where:

- `L0muon`: requires at least one muon with transverse momentum $p_T > 1.48$ GeV (in 2011) and $p_T > 1.76$ GeV (in 2012) .
- `Hlt1TrackAllL0`: requires a good quality track with $p_T > 1.6$ GeV and displaced from the primary vertex (PV).
- `Hlt1TrackMuon`: requires a good quality track with $p_T > 1$ GeV, displaced from primary vertex (PV) and passing the muon identification criteria.

	Decay	Number of events
(a)	$B^0 \to K^{*0}\mu^+\mu^-$ (phase-space)	5.5M (stripped)
(b)	$B^0 \to J/\psi K^{*0}$ (benchmark sample)	2M
(c)	$B^0 \to K^{*0}\mu^+\mu^-$ (benchmark sample)	1M
(c)	$\Lambda_b^0 \to \Lambda(1520)\mu^+\mu^-$	1M
(c)	$\Lambda_b^0 \to pK^-\mu^+\mu^-$	2M
(c)	$B_s^0 \to \phi\mu^+\mu^-$	0.6M
(c)	$B^+ \to K^+\mu^+\mu^-$	1M

Table 3.1: List of the simulated samples used in the analysis. Sample (a) is used for deriving the effiency correction: the number of events in this case corresponds to the size of the dataset after the stripping procedure. Sample (b) is used to derive data-simulation corrections. Samples (b-c) are used to assess exclusive backgrounds levels.

- `Hlt2Topo[2,3,4]BodyBBDT` or `Hlt2TopoMu[2,3,4]BodyBBDT`: they combine different variables *e.g.* the minimum transverse momentum of the tracks $p_T^{\min}$, invariant mass of the candidates , distance of closest approach (DOCA), using a boosted decision tree.
- `Hlt2SingleMuon`: has trigger `Hlt1TrackMuon` and additionally a track $\chi^2 < 2$, $p_T > 1.3$ GeV, impact parameter IP>0.5 mm and $\chi^2(\text{IP}) > 200$.
- `Hlt2DiMuonDetached`: the di-muon system has a mass below 1 GeV, a $p_T > 1.5$ GeV and the muons have $\chi^2(\text{IP}) > 9$ and a decay length significance (DLS) >7.

At all stages the offline-candidates are required to be triggered on signal (`TOS`), i.e. the trigger decision is due solely to the presence of the candidate in the event. No significant advantage was found by allowing candidates triggered independently from signal (`TIS`), or extra trigger lines.

3.3 Stripping

In order to reduce the size of the data sample used in the rest of the analysis, the events need to pass a set of very loose cuts. This step of the selection is called "Stripping" within the LHCb experiment, and might be common to many different analyse sharing common features. Candidates for the decays $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$ are required to pass the stripping requirements, where a good quality for the B^0, the di-muon and the K^{*0} verteces are required, as well as reasonable invariant mass window. The complete set of the stripping requirements are outlined in Table A.1.

Stage	Triggers
L0	`L0Muon`
HLT 1	`Hlt1TrackAllL0` or `Hlt1TrackMuon`
HLT 2	`Hlt2Topo[2,3,4]BodyBBDT`, `Hlt2TopoMu[2,3,4]BodyBBDT`, `Hlt2SingleMuon` or `Hlt2DiMuonDetached`

Table 3.2: List of triggers used in the analysis

3.4 Pre-selection

Candidates passing the stripping line are additionally required to satisfy other criteria, as cuts on the primary vertex position or hadrons particle identification (PID) requirements. The whole pre-selection is detailed in Table A.2. Differently from the stripping, which could be common to many channels, the pre-selection uses requirements more specific to the decay channel under study to further reduce the data sample to be analysed.

For example in the present case, additional requirements are made on the $K\pi$ invariant mass, $m_{K\pi}$. For the angular analysis, [1] an interval of $\pm 100\,\mathrm{MeV}/c^2$ around the nominal mass of the $K^*(892)^0$ is employed:

$$795.9 < m_{K\pi} < 995.9\,\mathrm{MeV}/c^2. \tag{3.1}$$

Actually, in order to correctly evaluate the contribution of the S-wave in the $K\pi$ system, a dedicated analysis has been done [2]. In this analysis, it has been shown that a wider $K\pi$ mass window is needed [6]:

$$633 < m_{K\pi} < 1200\ \ \mathrm{MeV}/c^2 \tag{3.2}$$

The multivariate selection described in Section 3.6 has been optimised using the largest window, so that it can be used for both the angular and the S-wave analyses.

3.5 Backgrounds from exclusive decays of b hadrons

This section describes how to reject the background from exclusive decays of b hadrons passing the trigger, stripping and preselection criteria. These backgrounds usually enter the signal region because a wrong mass hypothesis is assigned to one or more tracks or because one or more random tracks are added to a genuine B decay in the reconstruction. A special case is the charmonium resonances decaying in two muons, detailed in Sec. 3.5.1, which give exactly the same final state that the signal decay $B^0 \to K^{*0}\mu^+\mu^-$ but with a greatest decay rates.

The main strategy in order to reduce these backgrounds consists in using specific vetoes on some regions of the phase space. Simulated events have been used to estimate the efficiency of the pre-selection and of the specific veto cuts. The yields of each background have been estimated using the $b\bar{b}$ cross-section and the branching ratio of the different decays. The expected yields of each decay channel (in $3\,\mathrm{fb}^{-1}$) are summarised in Tables 3.3. The distributions from simulated events of invariant masses, angular variables and q^2 for each of the backgrounds and for the signal, is shown in Figure 3.1

More details on the different vetoes could be found in Ref. [7, 8].

3.5.1 Background from charmonia resonances

The tree-level decays $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ and $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^{*0}$, have exactly the same final state as $B^0 \to K^{*0}\mu^+\mu^-$ under study, though proceeding via charmonia resonances. Their contributions are large but can be suppressed requiring a cut on the di-muon invariant mass, q^2. The bin $q^2 \in [8.0, 11.0]\,\mathrm{GeV}^2/c^4$ (containing the $B^0 \to J/\psi\, K^{*0}$ resonant decay mode) and the bin $q^2 \in [12.5, 15.0]\,\mathrm{GeV}^2/c^4$ (containing the $B^0 \to \psi(2S)K^{*0}$ resonant decay mode) have been removed from the data samples before training the MVA. These vetoes have been used in Ref. [9] and are shown in Figure 3.2, where the contribution of the two channels is evident.

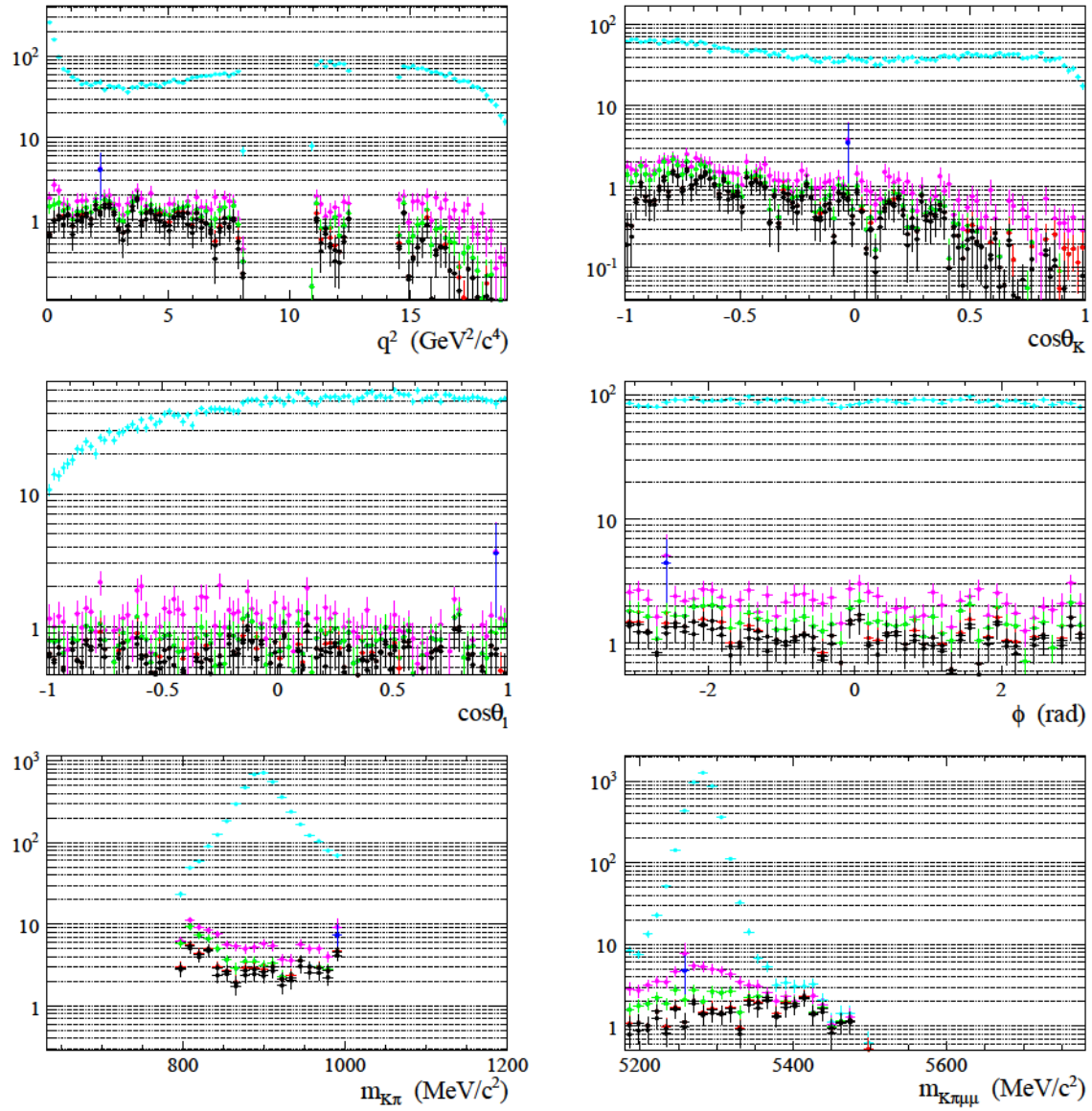


Figure 3.1: The distributions of the peaking backgrounds and of the signal in simulated samples is shown as stacked histograms, scaled to the expected number of candidates in 3 fb^{-1}. The distributions are: q^2 (top left), $\cos\theta_K$ (top right), $\cos\theta_l$ (middle left), ϕ (middle right), $m_{K\pi}$ (bottom left), and B candidate mass, $m_{K\pi\mu\mu}$ (bottom right). The channels shown are: $B^0 \to K^{*0}\mu^+\mu^-$ (signal) in **cyan**; signal mode with hadron swaps in **pink**; $B_s^0 \to \phi\mu\mu$ in **green**; $B^+ \to K^+\mu^+\mu^-$ in **red**; and samples of $\Lambda_b^0 \to pK^-\mu^+\mu^-$ and $\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-$ in black.

Channel	after preselection, before vetoes		after vetoes and selection	
	Estimated events	% signal	Estimated events	% signal
$\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-$	$(1.0\pm0.5)\times10^3$	19 ± 8	51 ± 25	1.0 ± 0.4
$\Lambda_b^0 \to pK^-\mu^+\mu^-$	$(1.0\pm0.5)\times10^2$	1.9 ± 0.8	5.7 ± 2.8	0.11 ± 0.05
$B^+ \to K^+\mu^+\mu^-$	28 ± 7	0.55 ± 0.06	1.6 ± 0.5	0.031 ± 0.006
$B_s^0 \to \phi\mu^+\mu^-$	$(3.2\pm1.3)\times10^2$	6.2 ± 2.1	17 ± 7	0.33 ± 0.12
signal swaps	$(3.6\pm0.9)\times10^2$	6.9 ± 0.6	33 ± 9	0.64 ± 0.06
$B^0 \to J/\psi K^{*0}$ swaps	$(1.3\pm0.4)\times10^2$	2.6 ± 0.4	2.7 ± 2.8	0.05 ± 0.05
$B^0 \to J/\psi K^{*0}$	70 ± 22	1.35 ± 0.28	59 ± 19	1.14 ± 0.26
$B^+ \to K^{*+}\mu^+\mu^-$	0	0	0	0

Table 3.3: Estimated yields, and percentage relative to estimated signal yield, of peaking background events before and after the vetoes. The dominant uncertainty contributing to these numbers is in $\sigma_{b\bar{b}}$ and the estimate of $\mathcal{B}\left[\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-\right]$.

3.5.2 The contribution of the resonances in the low q^2 region

The angular distribution of our signal can be modified by decays with the same final state but proceeding via intermediate resonances. The case of the two charmonium resonances, $J/\psi(1S)$ and $\psi(2S)$, has just been discussed previously (Section 3.5.1) and vetoes are imposed to remove their contributions. On the low q^2 region, some lighter di-muon resonances could also exist: ω, ρ and ϕ. The study of these low mass resonances is based on the previous study [10] of $B_d^0 \to K^{*0}e^+e^-$ decays at low q^2, and has been adapted for the $B^0 \to K^{*0}\mu^+\mu^-$ mode.

The contribution of the ω, ρ and ϕ resonances is assessed estimating the branching ratios of $B^0 \to K^{*0}V(V \to \mu^+\mu^-)$, (with V= ρ,ω,ϕ) and comparing it with the branching ratio of the non-resonant channel $B^0 \to K^{*0}\mu^+\mu^-$. In the following calculations, the branching ratios of V$\to e^+e^-$ are used, since they are known to a better precision than those for V$\to \mu^+\mu^-$, and a correction factor takes into account the difference between the electron and the muon channels. This correction factor is defined as:

$$(1+2\frac{m_\mu^2}{m_V^2})\sqrt{1-4\frac{m_\mu^2}{m_V^2}} \tag{3.3}$$

The correction turns out to be small and equal to 0.998 for both $\rho(770)$ and $\omega(782)$ and 0.9993 for ϕ.

The branching ratios of the resonant decays can then be computed as follows:

$$\begin{aligned}\mathcal{B}(B_d \to K^{*0}\rho(\rho \to \mu^+\mu^-)) &= \mathcal{B}(B_d \to K^{*0}\rho)\times\mathcal{B}(\rho \to e^+e^-)\times0.998\\ &= (3.9\pm1.3)\times10^{-6}\times(4.72\pm0.05)\times10^{-5}\times0.998\\ &= (1.84\pm0.61)\times10^{-10}\end{aligned}$$

$$\begin{aligned}\mathcal{B}(B_d \to K^{*0}\omega(\omega \to \mu^+\mu^-)) &= \mathcal{B}(B_d \to K^{*0}\omega)\times\mathcal{B}(\omega \to e^+e^-)\times0.998\\ &= (2.0\pm0.5)\times10^{-6}\times(7.28\pm0.14)\times10^{-5}\times0.998\\ &= (1.45\pm0.36)\times10^{-10}\end{aligned}$$

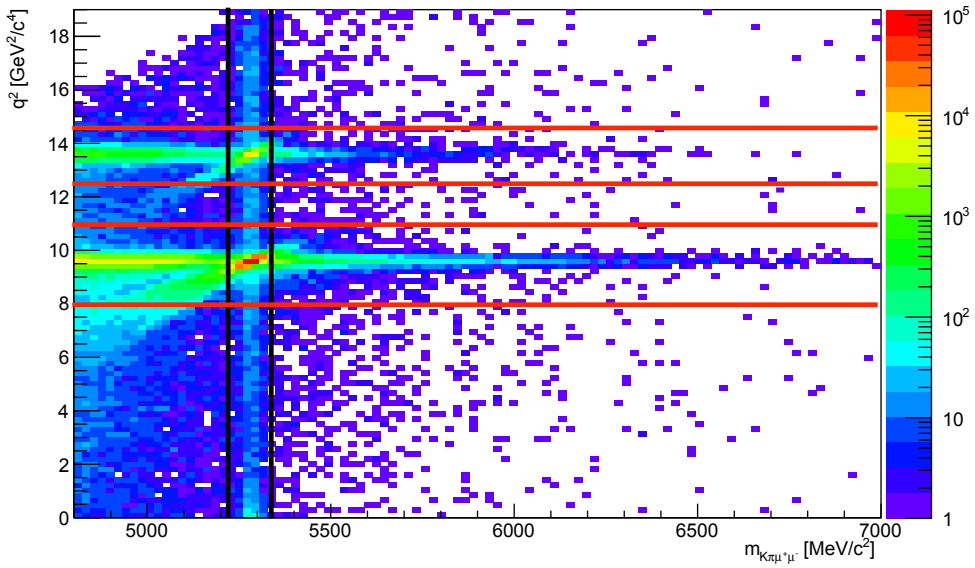


Figure 3.2: The $\mu^+\mu^-$ versus $K\pi\mu^+\mu^-$ invariant mass distribution of $B^0 \to K^{*0}\mu^+\mu^-$ candidates in data after pre-selection. The $J/\psi(1S)$ and $\psi(2S)$ vetoes are indicated by the red lines. The $B^0 \to K^{*0}\mu^+\mu^-$ decay is visible within the black lines.

$$\begin{aligned}\mathcal{B}(B_d \to K^{*0}\phi(\phi \to \mu^+\mu^-)) &= \mathcal{B}(B_d \to K^{*0}\phi) \times \mathcal{B}(\phi \to e^+e^-) \times 0.9993 \\ &= (1.00 \pm 0.05) \times 10^{-5} \times (2.954 \pm 0.03) \times 10^{-4} \times 0.9993 \\ &= (2.95 \pm 0.15) \times 10^{-9}\end{aligned}$$

For a precise estimation of the contributions the resonances to the non-resonant channel, one has to compare the branching ratios integrating them in the same interval, for example a 2Γ region around the mass of the resonance, where Γ is the natural width of the resonance. The branching ratio for the ρ resonance in that region is 1.55×10^{-10} while the branching ratio of $B^0 \to K^{*0}\mu^+\mu^-$ (estimated from the $B^0 \to K^{*0}e^+e^-$) is 1.5×10^{-7} within the same interval. This results in a ρ resonance contribution of 0.1% with respect to the non resonant channel. Similarly one can compute the ω resonance contribution, which is found to be 1.4% with respect to the non resonant channel. For the ϕ resonance the results turn out to be different. The resonant branching ratio in the $\pm 2\Gamma$ interval around the ϕ nominal mass is 2.4×10^{-9}, which is of the same order as the non resonant branching ratio. Therefore, the ϕ could be observable in our dataset.

The evaluation would not be complete if we were not taking into account the impact of the interference between the resonant channel and the non resonant one. This interference will depend on the relative phase of the two corresponding diagrams. Since this is unknown, we aim to estimate the maximum possible effect, which corresponds to a phase difference of $\pi/2$. In this most conservative scenario, it is the imaginary part B of the Breit-Wigner describing the resonance that gives the maximum interference. It can be evaluated from the following relation:

$$\int_{m_V-2\Gamma}^{m_V+2\Gamma} \left[\frac{A}{q} + \frac{B}{1+(q-m_V)^2/(\Gamma_V/2)^2}\right]^2 2qdq, \tag{3.4}$$

giving the interference term:

$$4AB\int_{m_V-2\Gamma_V}^{m_V+2\Gamma_V}\left[\frac{1}{1+(q-m_V)^2/(\Gamma_V/2)^2}\right]dq = 4AB\times\Gamma_V\times\arctan(4) \tag{3.5}$$

where, m_V and Γ_V are the mass and the width of the resonances, while the first term $\frac{A}{q}$ is the non resonant term.

The A term can be calculated using the value of the non resonant branching ratio in the 2Γ interval:

$$\begin{aligned}
\text{for } \rho: &\int_{m_\rho-2\Gamma_\rho}^{m_\rho+2\Gamma_\rho}\frac{A^2}{q^2}2qdq = 2A^2\ln\left(\frac{m_V+2\Gamma}{m_V-2\Gamma}\right) = 1.5\times10^{-7}\Rightarrow A = 3.0\times10^{-4}\\
\text{for } \omega: &\int_{m_\omega-2\Gamma_\omega}^{m_\omega+2\Gamma_\omega}\frac{A^2}{q^2}2qdq = 5.8\times10^{-9}\Rightarrow A = 2.55\times10^{-4}\\
\text{for } \phi: &\int_{m_\phi-2\Gamma_\phi}^{m_\phi+2\Gamma_\phi}\frac{A^2}{q^2}2qdq = 2.4\times10^{-9}\Rightarrow A = 2.68\times10^{-4}
\end{aligned}$$

The imaginary part of the Breit-Wigner, B, can be computed from the known branching fractions of the resonant channels $B_d \to K^{*0}V\times\mathcal{B}(V\to\mu^+\mu^-)$:

$$\begin{aligned}
\text{for } \rho: &\int_0^\infty\left[\frac{B}{1+(q-m_V)^2/(\Gamma_V/2)^2}\right]^2 2qdq = 1.84\times10^{-10}\Rightarrow B = 3.18\times10^{-8}\\
\text{for } \omega: &\int_0^\infty\left[\frac{B}{1+(q-m_V)^2/(\Gamma_V/2)^2}\right]^2 2qdq = 1.45\times10^{-10}\Rightarrow B = 1.18\times10^{-7}\\
\text{for } \phi: &\int_0^\infty\left[\frac{B}{1+(q-m_\phi)^2/(\Gamma_\phi/2)^2}\right]^2 2qdq = 2.95\times10^{-9}\Rightarrow B = 6.55\times10^{-7}
\end{aligned}$$

Finally, substituting A and B in Eq 3.5, we obtain the contribution of the interference in the conservative scenario. For the ρ it is 7.54×10^{-9}, which represents a contribution of 5% with respect to the non resonant channel. The same computation for the ω resonance gives an interference contribution of 1.36×10^{-9}, corresponding to 16% with respect to the non resonant channel. Even if more sizeable because of the interference effect, the contributions of these two resonances are still too small to be seen with the current dataset. For this reason they are safely neglected and no veto is applied to suppress them.

For the ϕ resonance, the non interfering term alone, gave a contribution from the ϕ at the level of 2.4×10^{-9}, so to avoid any risk of polluting the angular distribution with events coming from the $B_d \to K^{*0}\phi(\phi\to\mu^+\mu^-)$ decay, it has been decided to veto the ϕ region by removing the $q^2\in[0.98,1.1]\,\text{GeV}^2/c^4$ region.

Indeed, among the three resonances under study, only the ϕ resonance has been seen on data, as shown in Fig. 3.4

3.5.3 The contribution of $\Lambda_b^0 \to pK^-\mu^+\mu^-$

The peaking background from $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decays arise when the p is reconstructed as either of the hadron candidates. The dominant contribution to the $pK^-\mu^+\mu^-$ final state is expected to be from the $\Lambda^*(1520)^0 \to pK^-$ resonance. We veto these events by using an

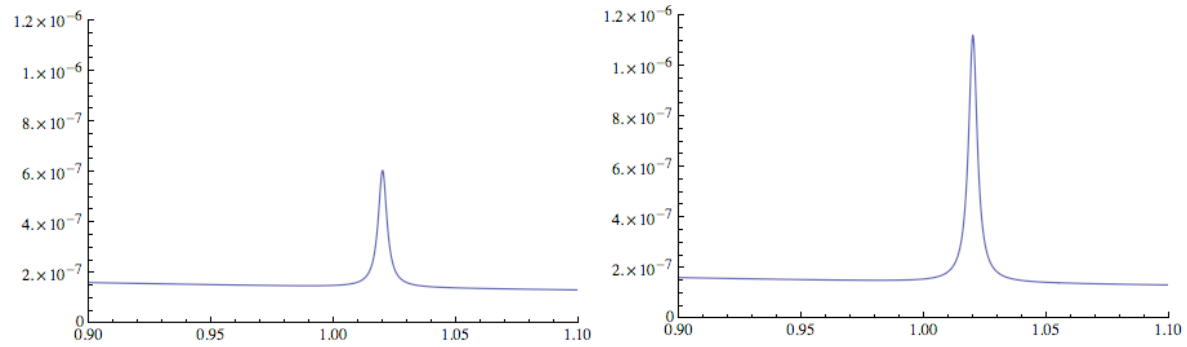


Figure 3.3: Computation of the branching ratio of the ϕ resonance without interference with the di-muon non resonant channel (left) and with a maximal interference (right). The interference can lead to triple the number of events expected from the decay involving the ϕ resonance.

alternative mass hypotheses and removing events in regions of mass and the particle information of the proton, $\mathrm{DLL}_{p\pi}$, as detailed below. The specific background can occur in two ways.

The simplest is when the p is reconstructed as the "π^+" candidate. In this case, changing the hypothesis of the π track, from the pion mass to the proton mass, and computing accordingly the invariant mass, $m_{(\pi\to p)K\mu\mu}$, should be consistent with Λ_b^0. The contribution from these events is removed by vetoing events with $m_{(\pi\to p)K\mu\mu}$ around the nominal Λ_b^0 mass with a proton-like $\mathrm{DLL}_{p\pi}$, *e.g.* candidates are removed if:

$$(5575 < m_{(\pi\to p)K\mu\mu} < 5665)\,\mathrm{MeV}/c^2 \tag{3.6}$$

$$\mathrm{DLL}_{p\pi}(\pi) > 0 \tag{3.7}$$

The second way $\Lambda_b^0 \to pK^-\mu^+\mu^-$ decays can contribute is when the p is reconstructed as the "K^+" candidate and the K^- is reconstructed as the "π^-". In this case, both hadron track mass hypotheses need to be changed. The resulting invariant mass, $m_{(K\to p)(\pi\to K)\mu\mu}$, should be consistent with Λ_b^0. This contribution is removed by a similar requirement on $m_{(K\to p)(\pi\to K)\mu\mu}$ and that the pion candidate has a kaon-like $\mathrm{DLL}_{K\pi}$. That is:

$$(5575 < m_{(K\to p)(\pi\to K)\mu\mu} < 5665)\,\mathrm{MeV}/c^2 \tag{3.8}$$

$$\mathrm{DLL}_{K\pi}(\pi) > 0 \tag{3.9}$$

3.5.4 Mis-identified $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$ decays

The decay $B^0 \to K^{*0}\mu^+\mu^-$ itself constitutes a peaking background if the K^+ is misidentified as a π^+ and the π^- is misidentified as a K^-. These misidentified candidates can be separated from correctly identified signal candidates since, if the π^- and K^+ mass assignments are swapped, the $K\pi$ invariant mass ($m_{K\leftrightarrow\pi}$) should then be consistent with that of a K^{*0}, so candidates are rejected if the kaon and pion satisfy:

$$\mathrm{DLL}_{K\pi}(K) + 10 < \mathrm{DLL}_{K\pi}(\pi) \tag{3.10}$$

$$(795 < m_{K\leftrightarrow\pi} < 995)\,\mathrm{MeV}/c^2. \tag{3.11}$$

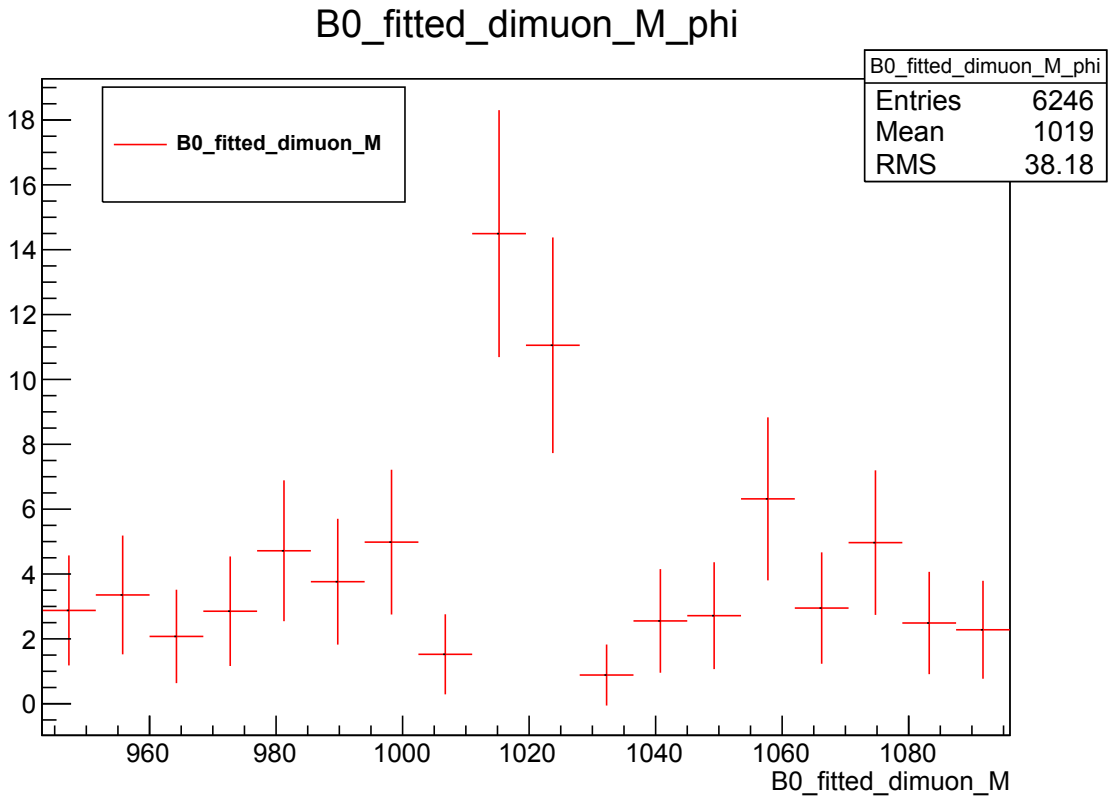


Figure 3.4: ϕ resonance seen with the 3fb^{-1} of data.

In addition, a requirement is made of the difference in the hadrons' $\text{DLL}_{K\pi}$ variables.

$$\text{DLL}_{K\pi}(K) - \text{DLL}_{K\pi}(\pi) > 0. \tag{3.12}$$

This removes remaining hadron PID swaps. The cut is usually referred to as "diagonal", forming a diagonal exclusion region in the $\text{DLL}_{K\pi}(K)$-$\text{DLL}_{K\pi}(\pi)$ plane, as shown in Figure 3.5.

Another way in which the $B^0 \to J/\psi K^{*0}$ decay contributes as a peaking background, is if the π^- (K^+) is misidentified as a μ^- (μ^+) and the μ^- (μ^+) is misidentified as a π^- (K^+). This is a so-called 'double-swap'. Candidates for which the π^- and μ^- are misidentified can be separated from signal candidates as, if the π^- and μ^- mass assignments are swapped, the $\pi\mu$ invariant mass $m_{(\pi\to\mu)\mu}$ should be consistent with that of a J/ψ. The equivalent quantity for candidates where the K^+ and μ^+ are misidentified, the $K\mu$ invariant mass $m_{(K\to\mu)\mu}$, is calculated by swapping the K^+ and μ^+ mass assignments. The candidates are rejected if

$$3036 < m_{(\pi\to\mu)\mu} < 3156\,\text{MeV}/c^2 \tag{3.13}$$

and the pion satisfies either the `IsMuon` criteria or $\text{DLL}_{\mu\pi} > 5.0$; or if

$$3036 < m_{(K\to\mu)\mu} < 3156\,\text{MeV}/c^2 \tag{3.14}$$

and the kaon satisfies either the `IsMuon` criteria or $\text{DLL}_{\mu\pi} > 5.0$.

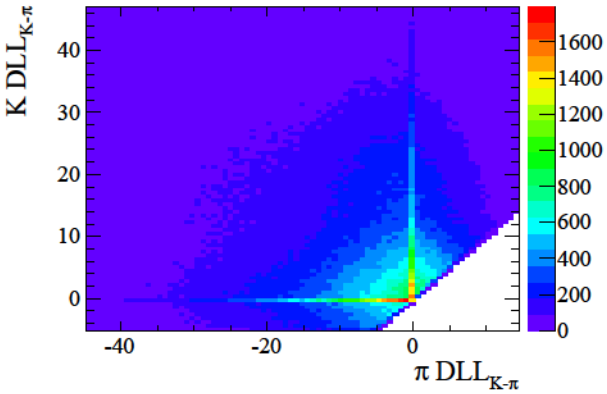


Figure 3.5: The effect of the "diagonal" cut. For the angular analysis the exclusion boundary is formed by the line $y = x$ in this plane. This is Eqn 3.12.

3.5.5 The contribution of $B_s^0 \to \phi\mu^+\mu^-$

The decay $B_s^0 \to \phi(\to K^+K^-)\mu^+\mu^-$ (hereafter denoted $B_s^0 \to \phi\mu^+\mu^-$) contributes to the peaking background if the K^- is misidentified as a π^-. If the reconstructed π^- is assigned the nominal mass of a K^-, the $K\pi$ invariant mass ($m_{(\pi\to K)K}$) should then be consistent with that of a ϕ. Similarly, the $K\pi\mu\mu$ invariant mass ($m_{(\pi\to K)K\mu\mu}$) should then be consistent with that of a B_s^0. Candidates are rejected if

$$(5321 < m_{(\pi\to K)K\mu\mu} < 5411)\,\mathrm{MeV}/c^2 \tag{3.15}$$

and either :

$$(1010 < m_{(\pi\to K)K} < 1030)\,\mathrm{MeV}/c^2 \tag{3.16}$$

and the pion satisfies $\mathrm{DLL}_{K\pi} > -10$, or

$$(1030 < m_{(\pi\to K)K} < 1075)\,\mathrm{MeV}/c^2 \tag{3.17}$$

and the pion satisfies $\mathrm{DLL}_{K\pi} > 10$. These two vetoes are defined in separate regions of $m_{(\pi\to K)K}$ to reduce the number of $B^0 \to K^{*0}\mu^+\mu^-$ decays that are removed. The effect is shown in Fig. 3.6.

3.5.6 The $B^+ \to K^+\mu^+\mu^-$ decay with an additional slow π^-

A background contribution can be formed if a π^- from elsewhere in the event is added to a genuine $B^+ \to K^+\mu^+\mu^-$ decay to form a four-track final state. Given that $B^+ \to K^+\mu^+\mu^-$ decays accumulate at the nominal B^0 mass, this background can only contribute to the upper $m_{K\pi\mu\mu}$ mass sideband. These candidates are typically asymmetric in $\cos\theta_K$, due to a momentum imbalance between the kaon and pion. This modifies the $\cos\theta_K$ distribution of candidates in the upper mass sideband, potentially causing a mis-measurement of the angular observables. The candidates are removed by requiring $m_{K\pi\mu\mu} > 5380\,\mathrm{MeV}/c^2$ and $(5220 < m_{K\mu\mu} < 5340)\,\mathrm{MeV}/c^2$, where $m_{K\mu\mu}$ is the $K\mu\mu$ invariant mass.

3.5.7 Other sources of peaking background

The decay $B^+ \to K^{*+}\mu^+\mu^-$ contributes to the background when a π^- from elsewhere in the event is added to the $K^{*+} \to K^+\pi^0$ decay, where the π^0 is not reconstructed, to form a four-track

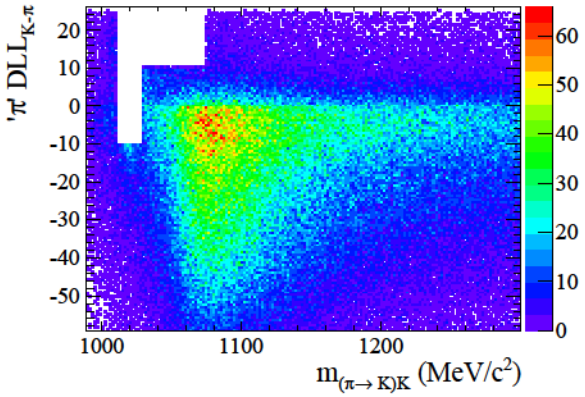


Figure 3.6: A plane showing the $B_s^0 \to \phi\mu^+\mu^-$ vetoed regions described in Section 3.5.5. This figure shows data candidates that fulfill Eqn 3.15, namely those candidates in the swapped hypothesis mass $m_{(\pi\to K)K\mu\mu}$ falling inside a window around the nominal B_s^0 mass. The excluded regions described in Eqns 3.16 and 3.17 are visible in the upper left-hand corner.

final state. These candidates cannot be distinguished from $B^0 \to K^{*0}\mu^+\mu^-$ decays by imposing requirements on invariant mass distributions or DLL quantities. Therefore, no veto criteria are imposed to remove this background. Its contribution is expected at the level of 1.5% of the signal yield accross the full $m_{K\pi\mu\mu}$ range and 0.5% within the signal mass window [7].

The $B^0 \to \rho^0(\to \pi^+\pi^-)\mu^+\mu^-$ (hereafter denoted $B^0 \to \rho^0\mu^+\mu^-$) decay contributes to the peaking background if the π^+ is misidentified as a K^+. The decay proceeds by the same process as the $B^0 \to K^{*0}\mu^+\mu^-$ decay, but it, is further suppressed by the ratio of CKM factors

$$\left|\frac{V_{cd}}{V_{cs}}\right|^2 \approx 0.05 \text{ and } \left|\frac{V_{td}}{V_{ts}}\right|^2 \approx 0.05 \tag{3.18}$$

for the resonant and non-resonant decays, respectively. Taking the rate of $\pi \to K$ mis-identification as 10% [11], the contribution of $B^0 \to \rho^0\mu^+\mu^-$ decays is calculated to be about 1% of the signal yield, and is therefore neglected.

3.6 Combinatorial background suppression by a multivariate selection

The combinatorial background originates from random combination of tracks in the event passing the trigger, stripping and preselection requirements and falling inside the invariant mass region under study. It is the most abundant background, but it can be highly reduced using the discriminating power of a number of variable exploiting the kinematics of the decay and the particle identification power of the detector. The simplest approach is to apply cuts on each of these variables one by one (rectangular cuts). But more powerful techniques exist, exploiting also the correlations between the variables. Indeed the most efficient way to use these variables is to combine them, making use of a multivariate approach (MVA) like, for example a Boosted Decision Tree (BDT). A visual comparison between the strategy of multivariate algorithms and rectangular cuts is shown in Figure 3.7.

The multivariate analysis presented here takes as a baseline the work performed for the previous LHCb analysis [12], and improves its performances while reducing the number of input

variables needed, in order to minimise the systematic uncertainties. The presentation of the classifier chosen for this analysis is given in Section 3.6.1. The input variables are described in section 3.6.3, while section 3.6.4 shows the comparison done among several MVAs in order to select the optimal.

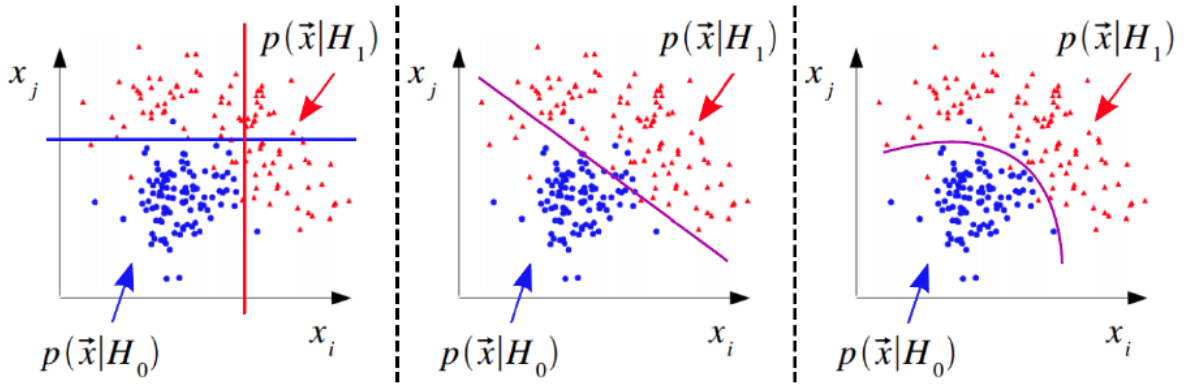


Figure 3.7: Illustration of (left) rectangular cut and (centre and right) multivariate algorithms, for two variable x_i and x_j. Blue and red points represent the signal and background samples.

3.6.1 The Boosted Decision Tree classifier

The Decision Tree [13] with Boosting technique [14] has been first used by the MiniBooNE collaboration [15]. The decision tree technique is a binary tree-structure, starting with a root node: a requirement that splits in two the variable space for both input samples. The input samples and each part of the variable space is then represented by a sub-node in the tree. One of the nodes is enriched with signal candidates while the other is enriched with background, as shown in Figure 3.8. During the "splitting" process, in each node, multiple requirements on each variable are tried, and one requirement on one variable is chosen to maximise the increase in separation between signal and background. This choice is made by comparing the G_{ini} index at the node to the sum of the G_{ini} indices at the two sub-nodes. The G_{ini} index is defined as :

$$G_{\text{ini}} = (\text{purity}) \times (1 - \text{purity}) = \frac{s}{s+b} \times \frac{b}{s+b}, \tag{3.19}$$

where s (b) is the amount of signal (background) at a given node.

The splitting process is repeated at each node until no improvement in separation is obtained or it stops once it has reached the minimum number of candidates in a "leaf" (node without splitting) or the tree depth, specified in the configuration.

Each leaf node represent a hypercube in the variable space and each of these leafs node is identified as "signal" or "background" depending on the class to which the majority of the events belongs. A background (signal) candidate in a leaf that the Decision Tree classifies as signal (background) is "misclassified".

A limitation of the Decision Tree is the fact that it is not robust against statistical fluctuations in the input data samples. A "boosting" procedure can be used to avoid this limitation. The boosting consists of a combination of several decision trees, to obtain a more performant classifier. Boosting is applied by giving a higher weight to misclassified events in the decision trees. At each iteration, misclassified events are weighted more heavily.

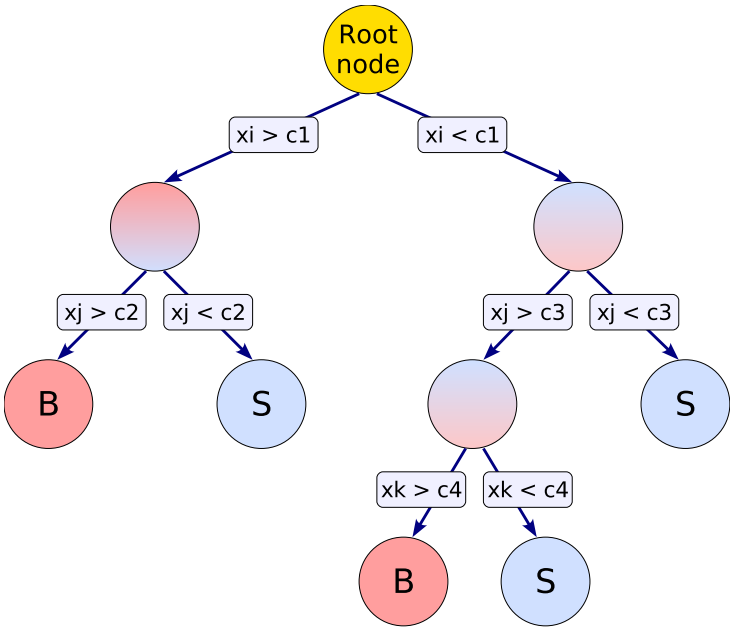


Figure 3.8: Illustration of a Decision Tree.

In the case of the Adaptive Boost (AdaBoost) [14], the boost weight, α_n, is calculated from the misclassification rate, $\varepsilon_n = \frac{\omega_{\rm mc}}{\omega_{\rm tot}}$, of the previous tree at the iteration n:

$$\alpha_n = \beta \log \frac{1-\varepsilon_n}{\varepsilon_n}, \tag{3.20}$$

where ω_{mc} is the sum-of-weights of all misclassified events, ω_{tot} is the sum-of-weights of all events, and β a tunable boosting-strength parameter.

The weights are then renormalised such as the sum-of-weights remains constant. Then the BDT response T(x) for the event x is defined as :

$$T(x) = \sum_{n=1}^{N} \alpha_n t_n(x), \tag{3.21}$$

where T(x) is a weighted sum of weak classifiers $t_n(x)$. The boosting adjusts the parameters to minimise the difference between the model response T(x) and the true value X, using a *loss function*, $L(T,X)$. In the AdaBoost case, the loss function is defined by an exponential:

$$L(T,X) = e^{-T(x)X} \tag{3.22}$$

For another boosting method used during this thesis, the Gradient Boost [16], the loss function is defined as binomial log-likelihood :

$$L(T,X) = \ln(1+e^{-2T(x)X}) \tag{3.23}$$

An important issue that needs to be taken care of is the risk of overtraining. It could happen that a multivariate classifier "learns too much", in the sense that it gets optimised on statistical fluctuations rather than on the general pattern of the data. In order to check overtraining, the

BDT input samples are split into two sub-samples. One is used to train the BDT, referred to as "training sample". The second one, called "test sample", is not used for the training and the BDT response is just applied to it to test the results. A comparison of the final distributions for signal and background, obtained in the test and training samples is made to check if the procedure was overtrained. An example is shown in Figure 3.9.

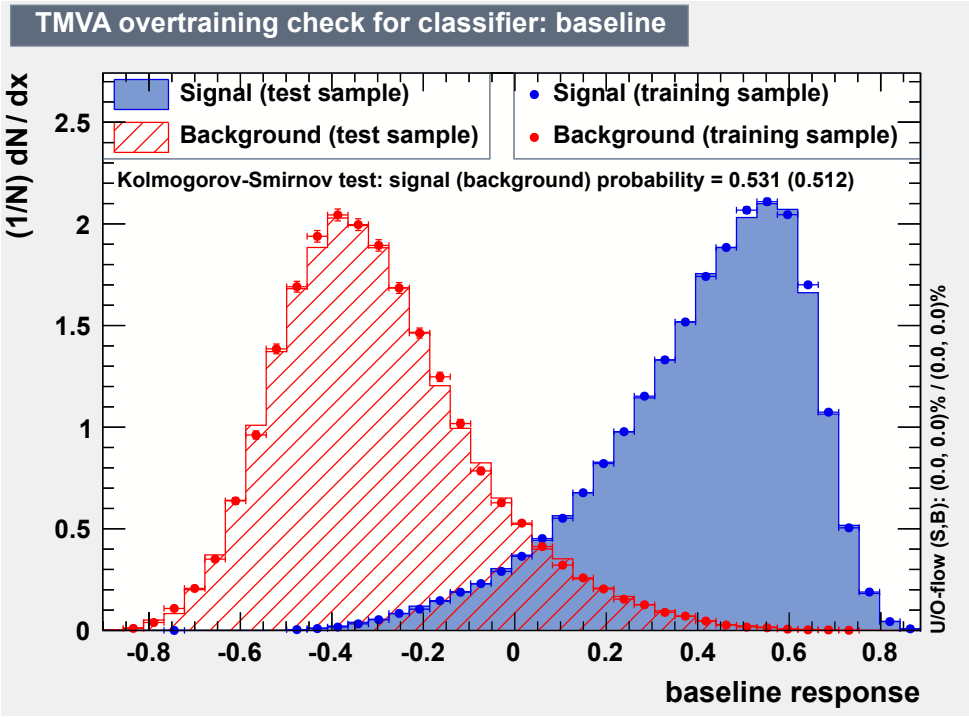


Figure 3.9: Example of a check for overtraining for the signal (blue) and background (red) samples. The good agreement between test and training sample and the result of the Kolmogorov-Smirnov test guarantee that there was no overtraining.

3.6.2 Selection of input samples

As previously mentioned, the MVA method needs calibration samples in order to be trained. For the signal sample we use data candidates from our control channel, the $B^0 \to J/\psi K^{*0}$ decay, described in Sec.3.6.2.1. For the background sample we use the upper mass sideband in our $B^0 \to K^{*0}\mu^+\mu^-$ data, described in Sec. 3.6.2.2

3.6.2.1 Signal proxy: the $s-weighted$ $B^0 \to J/\psi K^{*0}$ candidates

The *sPlot* technique [17] is employed to weight stripped $B^0 \to J/\psi K^{*0}$ candidates. This procedure statistically removes the background contribution such that the $B^0 \to J/\psi K^{*0}$ data sample can be considered to be “pure” signal. The result of the *sPlot* technique, after the fit to the un-weighted $B^0 \to J/\psi K^{*0}$ candidates, can be seen in Fig. 3.10, where a clear peak shows up.

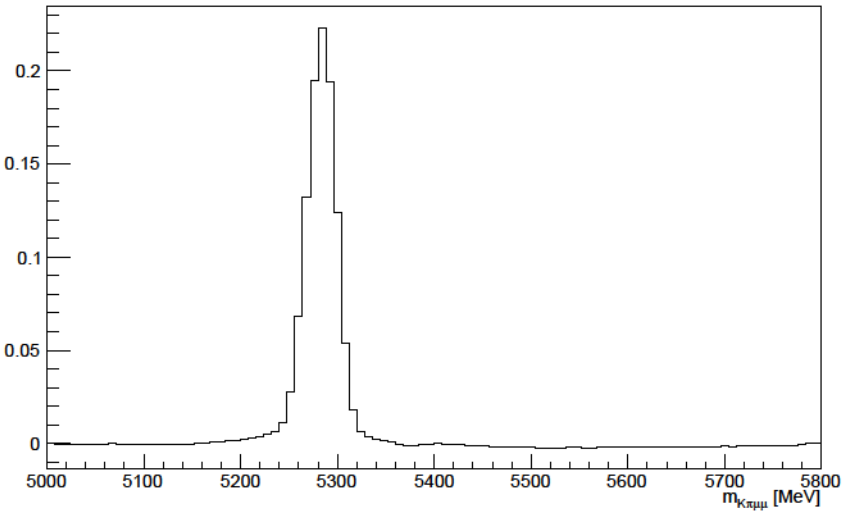


Figure 3.10: Invariant mass of the $B^0 \to J/\psi K^{*0}$ candidates with sWeight applied.

3.6.2.2 Background proxy: events in the upper mass sideband

The combinatorial background sample is extracted directly from the $B^0 \to K^{*0}\mu^+\mu^-$ data. On the left and right side of the signal mass window, one can define two sidebands composed exclusively of background events. Events from partially reconstructed B^0 mesons populate the low mass sideband. This removes the possibility to use this part as a calibration sample for the combinatorial background. Therefore, only events in the upper mass sideband are used as a proxy for the combinatorial background in the MVA training. The mass range used is $m_{K\pi\mu\mu} \in [5350, 7000]\,\text{MeV}/c^2$, as shown in the Figure 3.11. A comparison of the input variables distributions, between the smaller upper mass sideband ($m_{K\pi\mu\mu} \in [5350, 5780]\,\text{MeV}/c^2$), used in the previous analysis [4], and the present one has been performed. No differences have been found, as shown in Figure 3.12. The new upper mass sideband constitutes an improvement with respect to the previous analysis as it allows to almost double the number of background events available for the MVA training.

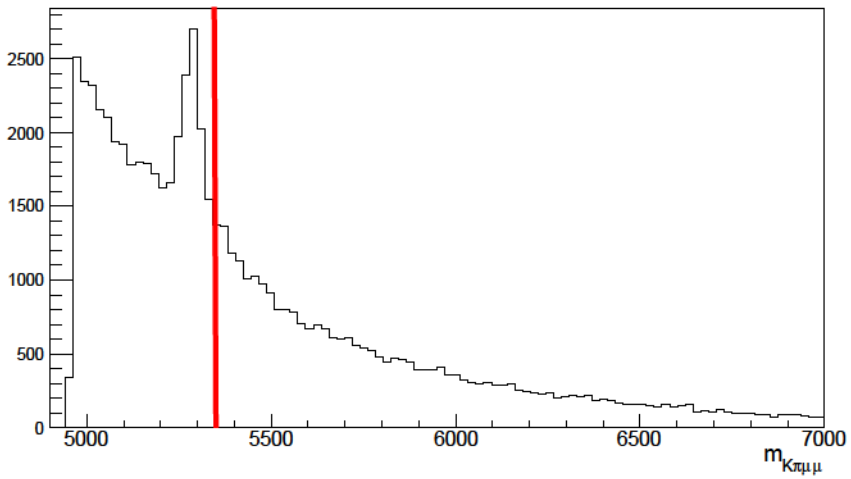


Figure 3.11: Invariant mass of the $B^0 \to K^{*0}\mu^+\mu^-$ candidates with the cut at 5350 MeV/c^2 defining the lower bound of the upper mass sideband.

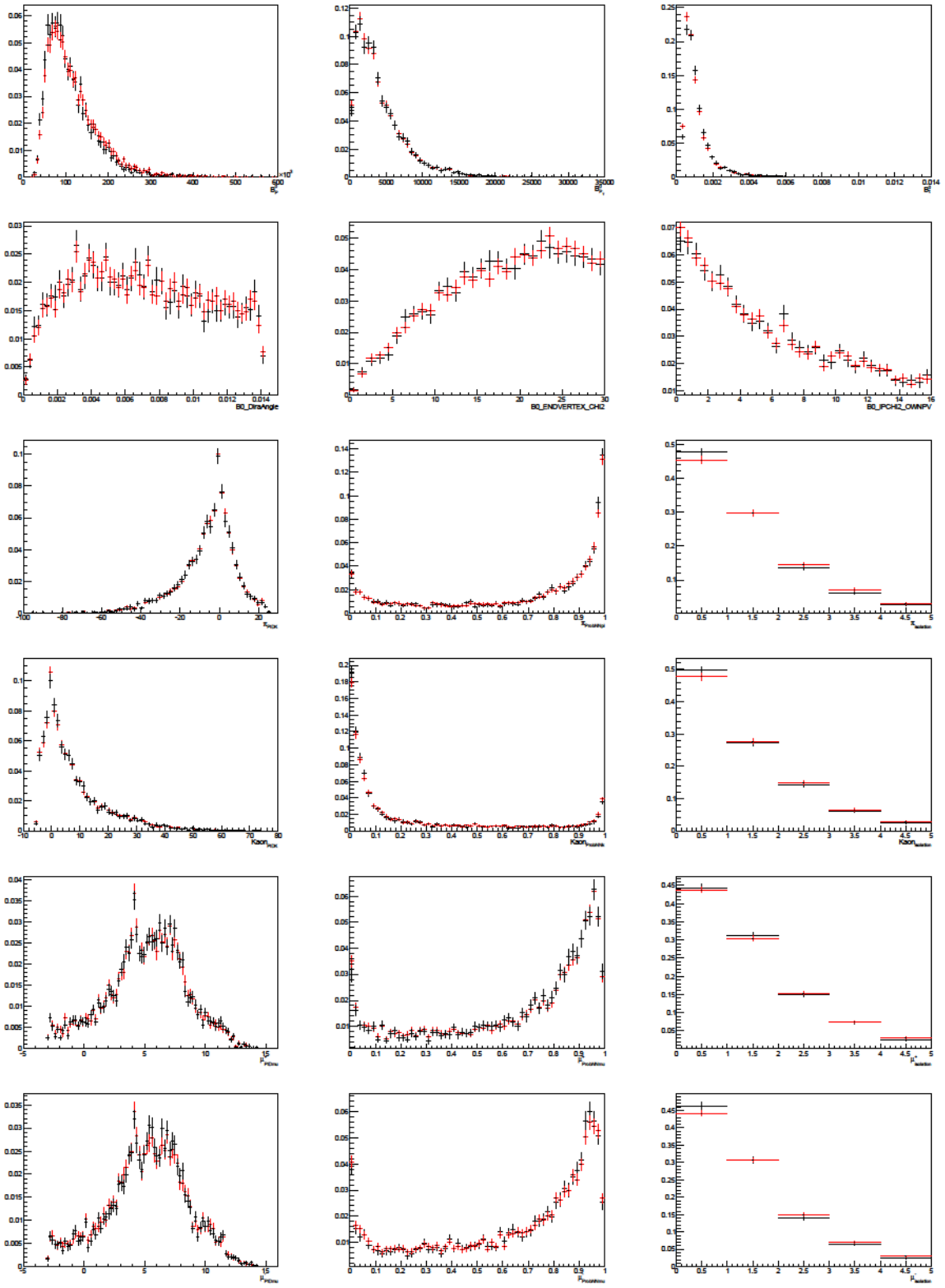


Figure 3.12: Distributions of several input variables for the old upper mass sideband window $m_{K\pi\mu\mu} \in [5350, 5780]\,\mathrm{MeV}/c^2$ (in black), and the new mass window $m_{K\pi\mu\mu} \in [5350, 7000]\,\mathrm{MeV}/c^2$ (in red). The distributions are in good agreement.

3.6.3 Discriminating variables

The choice of the discriminating variables used as input for the BDT takes as a starting point the list of those adopted in the previous analysis [4]. In order to minimise the systematic uncertainties, we choose to keep only some of them. We selected the variables having a reasonable agreement between data and simulation, the highest discriminating power and the lower correlation among them.

First of all we selected a set of 5 kinematic/topological variables and 2 PID variables. The kinematic/topological variables are :

- the B^0 lifetime
- the $K\pi\mu\mu$ vertex quality (χ^2)
- the momentum of the B^0 meson
- the transverse momentum of the B^0 meson
- the cosine of the angle between the momentum and the direction of flight from the primary to the decay vertices (DIRA).

The PID variables are:

- the $\mathrm{DLL}_{K\pi}$ for the pion and the kaon
- the $\mathrm{DLL}_{\mu\pi}$ for the muon tracks

The correlation between the different input variables is shown in Fig. 3.13, and can be seen to be rather low.

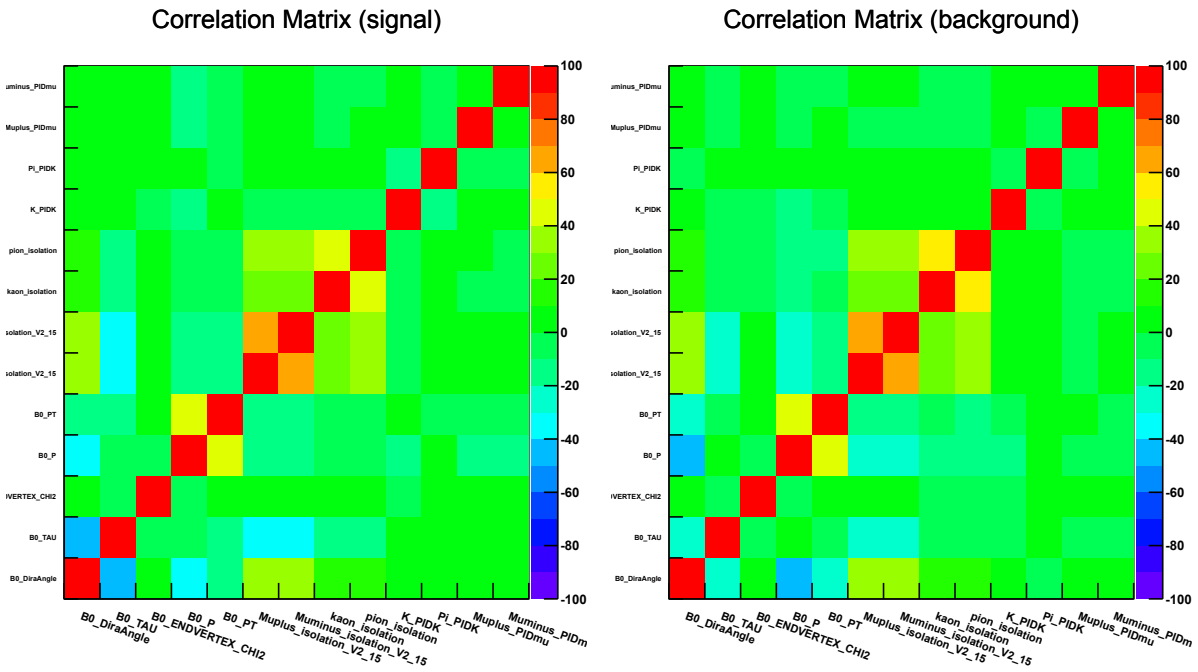


Figure 3.13: Correlation matrices of the input variables used in the BDT training for the signal (left) and background (right) samples.

This preliminary set of input variables was further improved, first introducing the isolation variables, described in section 3.6.3.1, and then using the Neural Network based PID variables, named "ProbNN", described in section 3.6.3.2.

3.6.3.1 Isolation variables

The isolation variables definition is based on the isolation used in the $B_s^0 \to \mu^+\mu^-$ analysis [18] and were used for $B^0 \to K^{*0}\mu^+\mu^-$ decays for the first time. More details about the optimisation of these isolation variables can be seen in Ref. [19].

We call **non-isolating tracks** the tracks in the event, other than those belonging to the selected $B^0 \to K^{*0}\mu^+\mu^-$ candidate, that can form a vertex with a signal candidate track. We denote in the following as **track** one of the tracks belonging to the signal candidate (designated by h in Fig. 3.14). The isolation variables is built counting how many **non-isolating tracks** in the event satisfy all the followings conditions, based on the variable described in Fig. 3.14:

- Minimum distance between the **non-isolating track** and the PV : **pvdist** $\in [0.5, 40]$ mm
- Minimum distance between **non-isolating track** and the $B^0 \to K^{*0}\mu^+\mu^-$ vertex : **svdist** $\in [-0.15, 30]$ mm
- DOCA between the **track** and the **non-isolating track** : DOCA < 0.13
- Track Impact Parameter (IP) significance : $ips > 3$
- Angle between the **track** and the **non-isolating track** : $\beta < 0.27$
- The quantity

$$f_c = \frac{|\overrightarrow{P}_h + \overrightarrow{P}_{trk}|\alpha^{h+\mathrm{trk},PV}}{|\overrightarrow{P}_h + \overrightarrow{P}_{trk}|\alpha^{h+\mathrm{trk},PV} + P_{\mathrm{T},h} + P_{\mathrm{T},trk}} < 0.6, \tag{3.24}$$

  where $\alpha^{h+\mathrm{trk},PV}$ is the angle, in radian, between the **track** and the **non-isolating track** candidate, $P_{\mathrm{T},h}$ and $P_{\mathrm{T},trk}$ are the transverse momenta with respect to the beam line.

The distribution of the isolation variable when the track h is the pion, the kaon and the muons can be seen in Figure 3.15.

For the specific case of the muons, the isolation variable was further optimized employing a multivariate approach. To do this the quantities described above, for each track in the event, are fed into a MVA, together with the three following variables:

- Track IP
- Track p_{T}
- Track χ^2/ndof

For training the MVA a simulated signal sample with a phase space model and an inclusive di-muon background sample were used. The best performance was obtained using a BDT classifier, which is henceforth referred to as BDTiso. The distribution of the new muons isolation variable is shown in Figure 3.16

The performance of a BDT can be represented in a plane showing the background rejection vs. the signal efficiency for various cut on the MVA classifier. Such a plot is called a Receiver Operating Characteristic (ROC) diagram, and can be used to compare the performances of different BDTs.

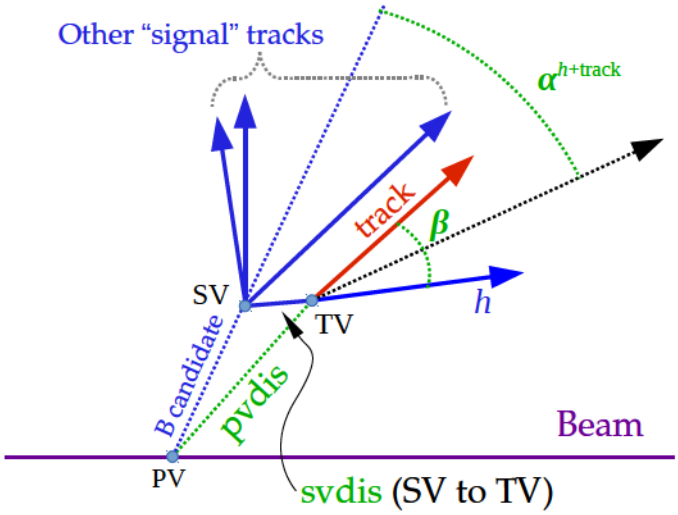


Figure 3.14: A diagram showing the variables used in the isolation variable of the track "h". Description is given in Section 3.6.3.1.

Different sets of isolation variables were added to the kinematic variables for training different BDTs and test their performances. The ROC diagram in Figure 3.17 shows that adding the isolation variables clearly improves the separation between signal and background comparing to using only the kinematic variables.

3.6.3.2 Particle Identification variables

In the previous analysis [4], among the particle identification variables available, only the one for the hadrons was used as input in the MVA. This variable is defined as:

$$\text{PIDK} = \Delta LL_{K/\pi} = \ln(L(K)/L(\pi)), \tag{3.25}$$

were $L(K)$ and $L(\pi)$ are the products of the likelihoods provided by the different sub-detectors: $L(h) = L(h)^{RICH} \times L(\text{non e})^{CALO} \times L(\text{non } \mu)^{MUON}$, with h being a kaon or a pion.

In the same way one can actually build the muon particle identification:

$$\text{PIDmu} = \Delta LL_{\mu/\pi} = \ln(L(\mu)/L(\pi)), \tag{3.26}$$

with $L(\mu) = L(\mu)^{\text{RICH}} \times L(\text{non e})^{\text{CALO}} \times L(\mu)^{\text{MUON}}$. The distribution of the PID variables, can be seen in Figure 3.18.

In this analysis the PIDmu variable, as well as new hadron particle identification variables, called the ProbNN, were tested as input variable in the MVA. The ProbNN variables are the result of a combination of the ΔLL variables in an artificial neural network (ANN), making use of additional information about the tracks, such as the P_T, the χ^2/ndf, as well as the DLLe, the DLLμ and other inputs from the sub-detectors (RICH, ECAL and HCAL). The distributions of the ProbNN variables are shown in Figure 3.19.

The comparison between the different sets of PID and ProbNN variables is shown in the ROC curve of Figure 3.20. Despite the better performance of the ProbNN variables, we chose to use the PID variables as input for our MVA, since at that time the ProbNN distributions were

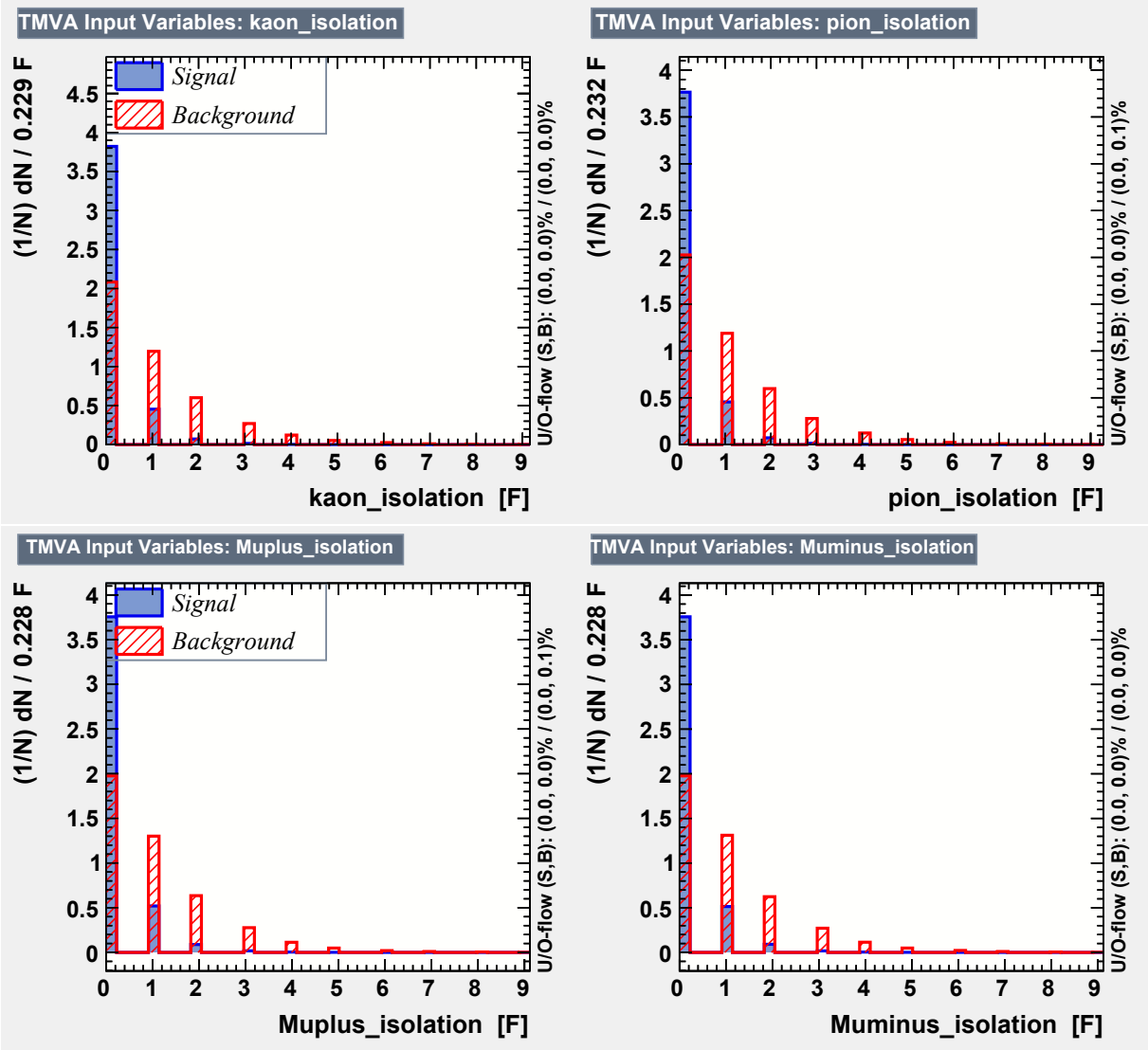


Figure 3.15: Distribution of the kaon and pion (top) and muons (bottom) isolation for signal (in blue) and background (in red) events.

not very well reproduced in the simulation, compared to the PID variables, and this would have led to a larger systematic uncertainty.

3.6.4 k-Folding of the data sample

When adopting an MVA approach, it is not possible to use in the analysis the same data sample used for the training procedure, as there is the risk to introduce a bias in the measurement. In order to be able to use the full data sample recorded by LHCb in 2011 and 2012, the k-Folding technique was used [20].

First of all, to avoid any bias due to the variations in the detector running conditions (e.g. different running energy, magnet polarities) the ordering of the events in the dataset is randomised; the data is then divided in $k = 10$ samples of equal sizes, both for the full background

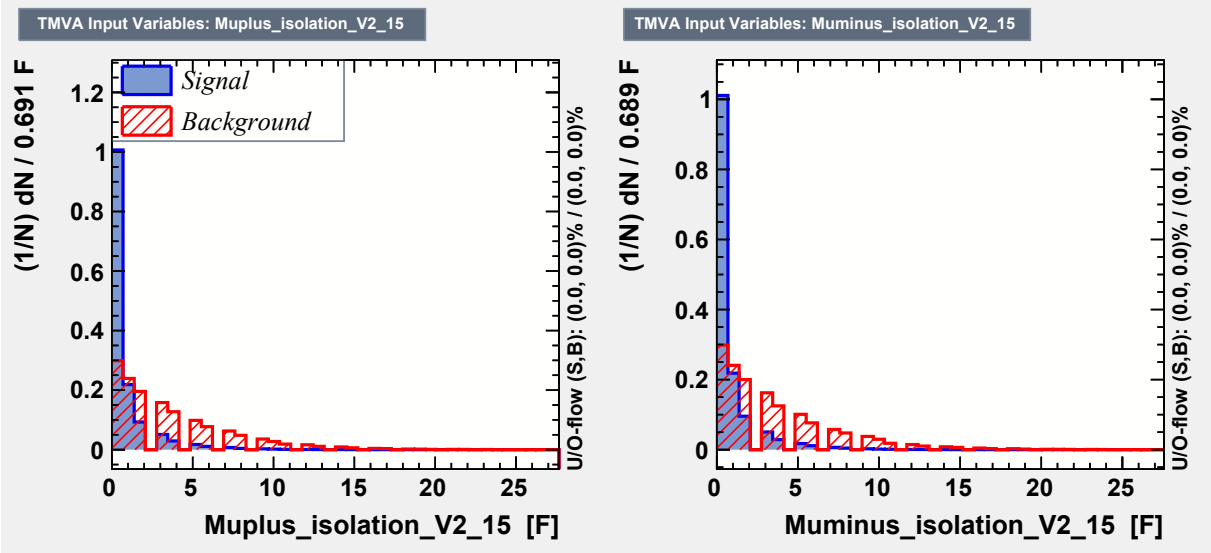


Figure 3.16: Distribution of muon isolation variables for signal (in blue) and background (in red) events.

sample, $B = \bigcup_i B^i$, and for the full *sWeighted* $B^0 \to J/\psi K^{*0}$ sample used as a signal proxy, $S = \bigcup_i S^i$.

Then, ten classifiers are trained, each using nine signal and nine background samples. More formally, the i^{th} classifier uses $\bigcup_{j,j\neq i} B^j$ and $\bigcup_{j,j\neq i} S^j$ training samples. The obtained i^{th} classifier is then applied to the i^{th} signal (S^i) and background (B^i) samples that were omitted from its training, as illustrated in Figure 3.21.

Such a training has several advantages compared to a standard procedure where the data sample is divided in two halfs only:

- The i^{th} and the j^{th} classifier have $\frac{8}{9}$ common events, which makes the classifiers responses more similar, reducing the systematic uncertainty (see Ref. [20] for a mathematical proof)
- The training samples are increased in size, as the 10-Folding technique allows to use 90% of the data for the training, compared to 50% in the standard case, which also leads to a better optimisation of the classifier.

After the training, each fold has its own BDT response. All the BDTs have similar responses, as expected and desired. After applying the different BDTs on the different datasets the BDT output values assigned to all the candidates in the all dataset are treated in the same way,*i.e.* a unique and identical cut on all the BDT responses is applied. The comparison of the ten classifiers responses is shown on Figure 3.22.

For simulation samples, which anyway were not used for training, the following quantity is used as a BDT estimator:

$$\mathrm{BDT_{MC}} = \frac{\sum_{\mathrm{i}=1}^{10} \mathrm{BDT_{Fold\ i}}}{10}, \tag{3.27}$$

where $\mathrm{BDT_{MC}}$ is the response for a simulated event, and $\mathrm{BDT_{Fold\ i}}$ is the response for the simulated event using the $\mathrm{i^{th}}$ fold's classifier.

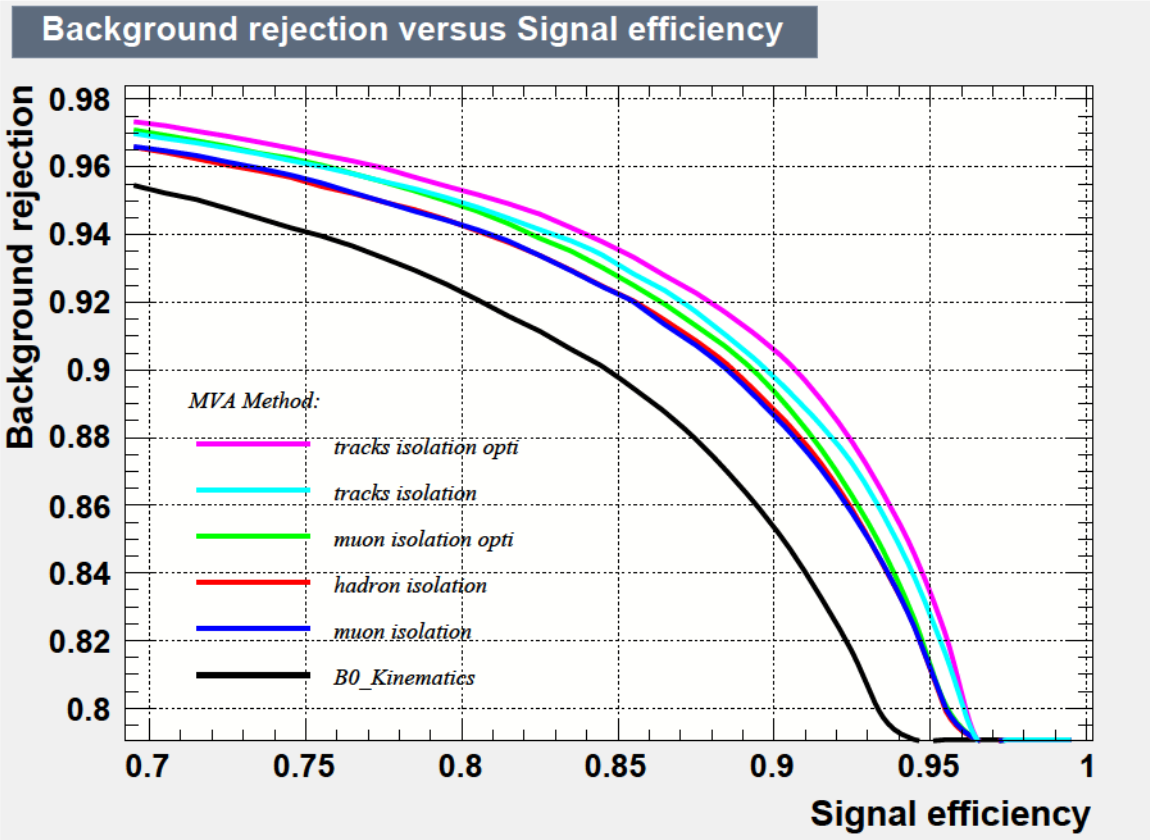


Figure 3.17: ROC curves for different set of input variables: 5 kinematic variables from B^0 meson (in black), 5 B^0 variables and hadron isolation (in red), 5 B^0 variables and muon isolation (in blue), 5 B^0 variables and the muon isolation variable (in green), 5 B^0 variables with the 4 final state particle isolation variables (in cyan), and 5 B^0 variables with the 4 final state particle optimised isolation variables (in purple),.

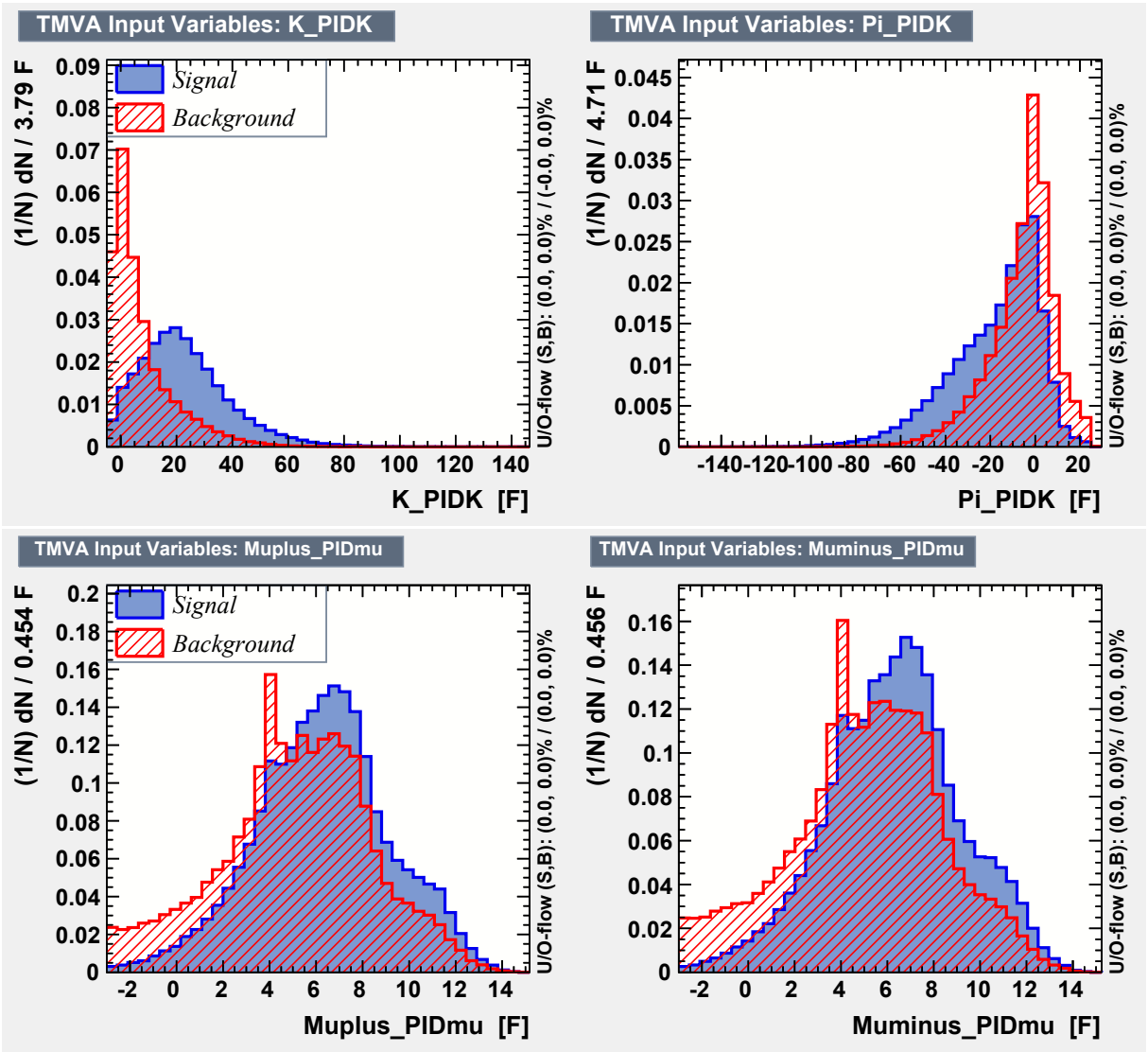


Figure 3.18: Distribution of hadron (top) and muon (bottom) PID variable for signal (in blue) and background (in red) events.

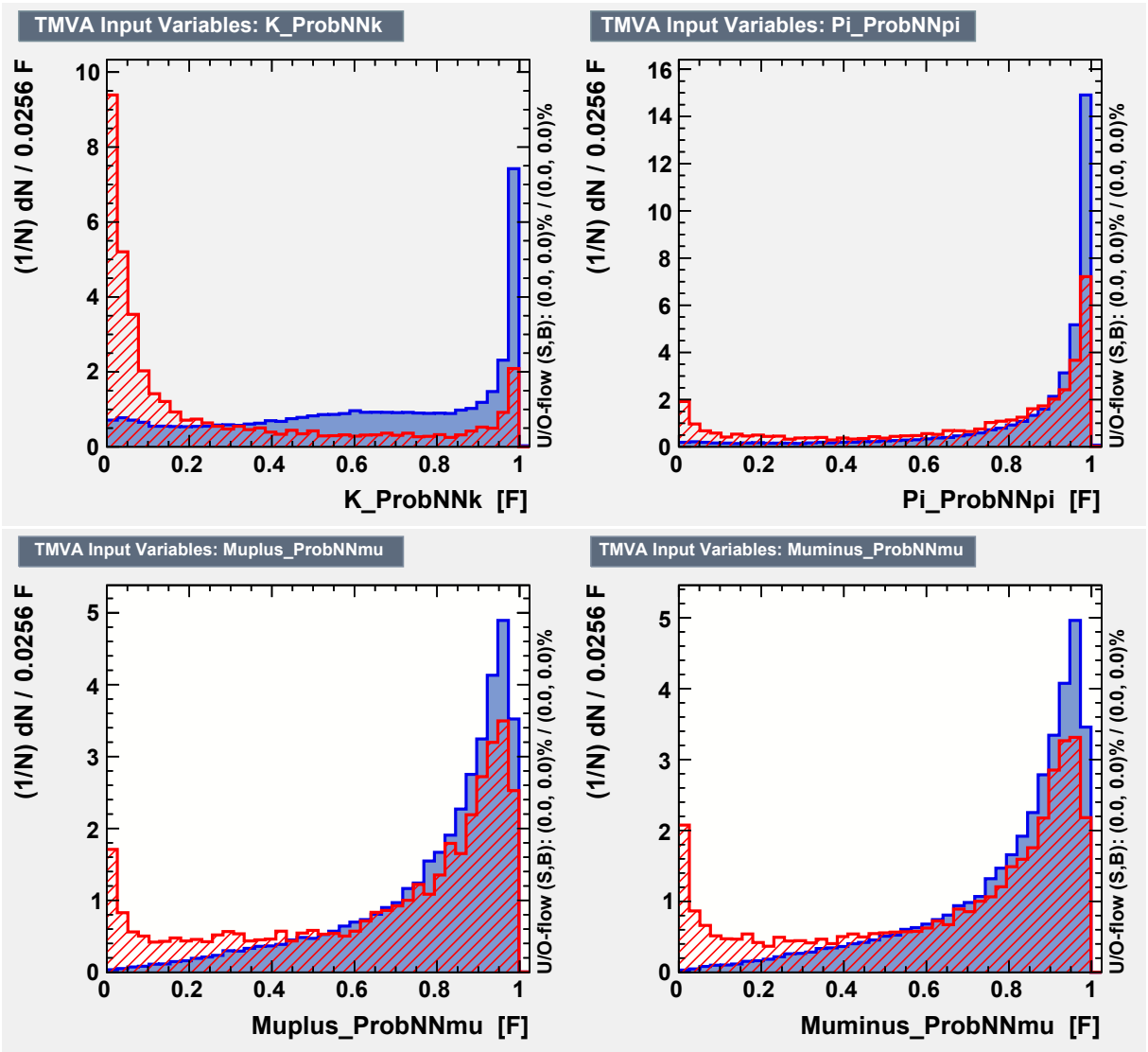


Figure 3.19: Distribution of hadron (top) and muon (bottom) ProbNN variable for signal (in blue) and background (in red) events.

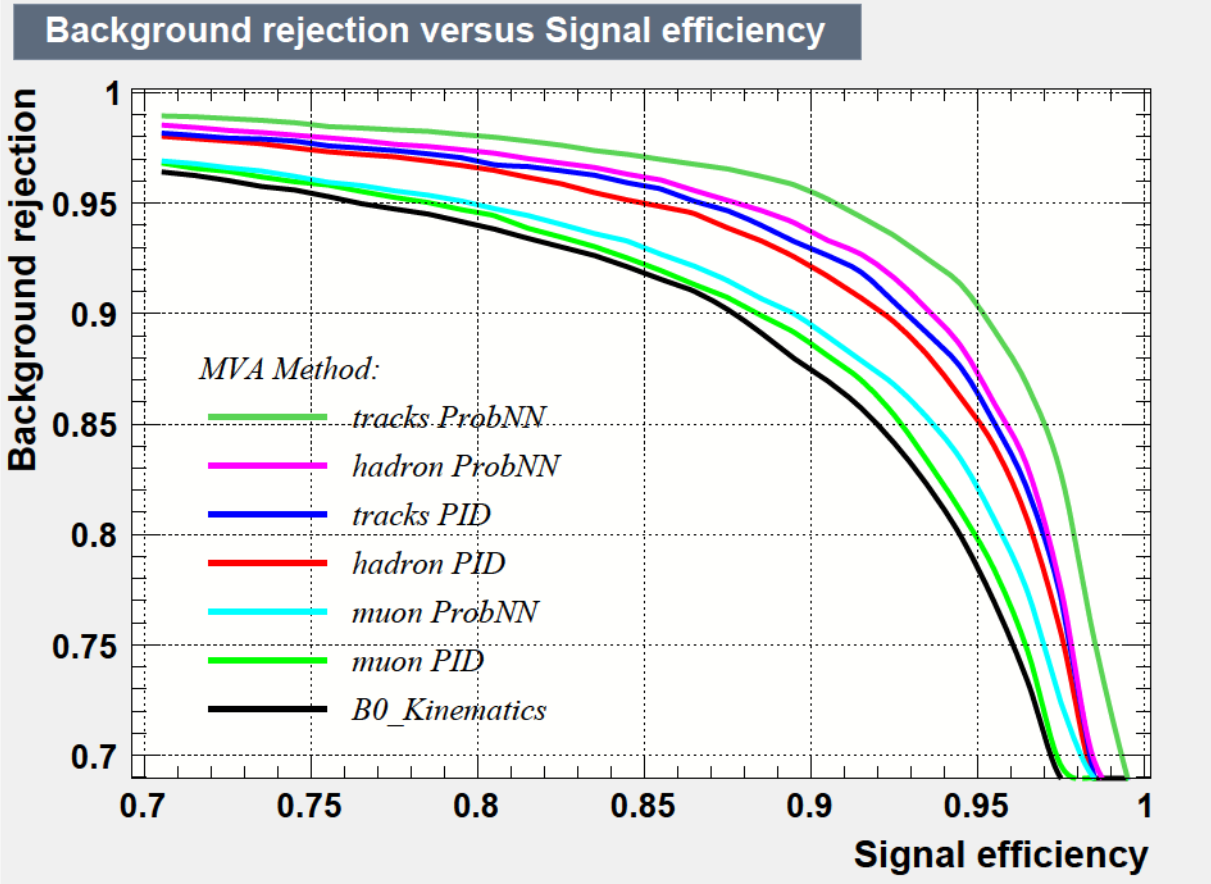


Figure 3.20: ROC curve for different sets of input variables. In black, the result for a BDT using 5 kinematic variable from B^0 meson, which is the baseline for assign the impact of adding the PID variables. The other curves shows the improvement on top of the baseline when adding: the hadron PID (in red); the muon PID (in green); the 4 final state particles PIDs (in blue); the hadron ProbNN (in purple); the muon ProbNN (in cyan); the 4 final state particles ProbNNs (in light green).

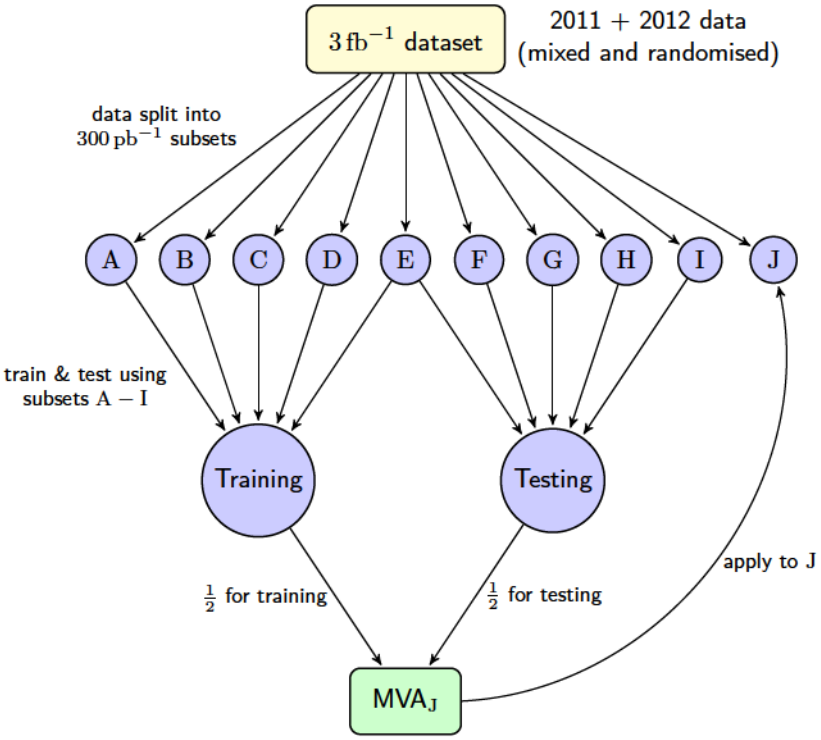


Figure 3.21: A schematic diagram of the 10-Folding technique. Arrows indicate the direction of the data "flow". The data is split randomly into 10 sub-datasets (A–J). The MVA for sample J (MVA$_J$) is trained and tested on samples A–K, then this response is evaluated on sample J only. This way, 90% of the dataset is used for every BDT training in an unbiased way.

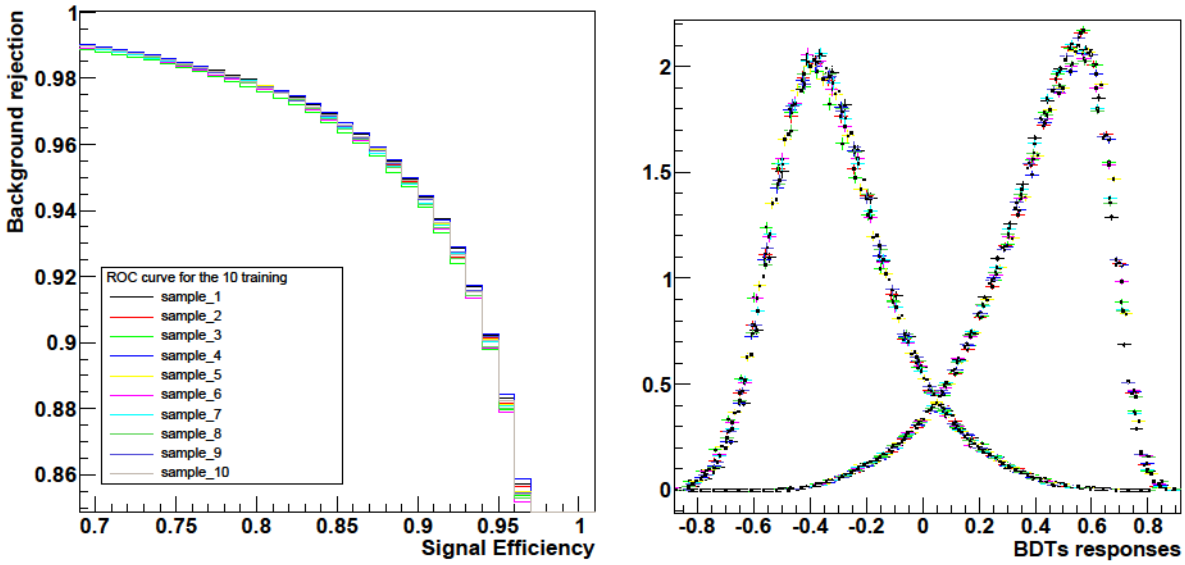


Figure 3.22: Performances of the ten classifiers (left) and the 10 classifiers response distributions (right), the signal (background) events are in the right (left) peak.

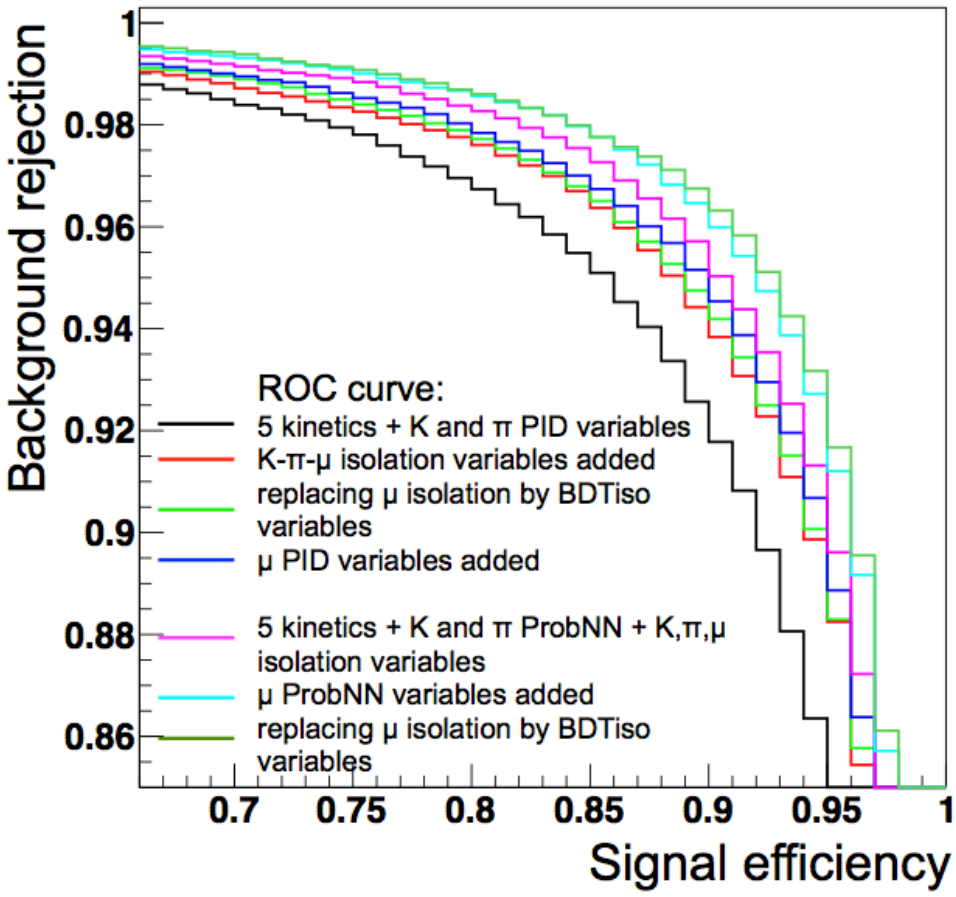


Figure 3.23: ROC curves of BDTs using different input variables. In black, the BDT trained with the 5 kinematic variables of the B^0 meson and the PID information for the kaon and the pion. In red (light green) the isolations variables (in particular BDTiso for the muon isolation) of the final state particle are added to the training. In dark blue the PID information of the muons are added. In pink the training uses the 5 kinematic variables of the B^0 meson and the ProbNN information for the kaon and the pion and the isolations variables. In cyan the ProbNN information of the muons are added to the BDT training. In green the muon isolation variables are replaced by the BDTiso isolation variables.

3.6.5 Training results

A summary of the different BDTs trained with the 10-folding technique and the different sets of input variables can be seen in Figure 3.23. The final choice, considering all the criteria discussed in the previous sections is the one for a BDT containing the following variables:

- the B^0 lifetime
- the $K\pi\mu\mu$ vertex quality (χ^2)
- the momentum and transverse momentum of the B^0 meson
- the cosine of the angle between the momentum and the direction of flight from the primary to the decay vertices.
- the $\mathrm{DLL}_{K\pi}$ for pion and kaon
- the $\mathrm{DLL}_{\mu\pi}$ for the muons
- the isolation variables of the pion, the kaon and the muons

and can be seen in the Figure. 3.23 with the dark blue line.

The cut on the BDT has been optimized using the following figure of merit : $\frac{\mathrm{nSig}_{\mu\mu}^{\mathrm{BDT}}}{\sqrt{\mathrm{nSig}_{\mu\mu}^{\mathrm{BDT}}+\mathrm{nBkg}_{\mu\mu}^{\mathrm{BDT}}}}$.

For each BDT cut, the signal yield in a window $m_{K\pi\mu\mu} \in (5230, 5330)$ is estimated by fitting $B^0 \to J/\psi K^{*0}$ events and scaling the yield obtained by the ratio of the total selection efficiency (apart from the BDT cut) obtained from simulation, and the ratio of branching fractions obtained from the PDG, between $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$:

$$\mathrm{nSig}_{\mu\mu}^{\mathrm{BDT}} = \mathrm{nSig}_{\mathrm{J}/\psi}^{\mathrm{BDT}} \times \frac{\epsilon_{\mu\mu}^{\mathrm{MC}} \times \mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)}{\epsilon_{J/\psi}^{\mathrm{MC}} \times \mathcal{B}(B^0 \to J/\psi K^{*0}) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)}$$

The background yield is estimated by fitting a part of the lower and upper mass sideband regions of $B^0 \to K^{*0}\mu^+\mu^-$ events and extrapolating the yield into the signal region ($\mathrm{nBkg}_{\mu\mu}^{\mathrm{BDT}}$). The upper mass sideband is defined as $m_{K\pi\mu\mu} \in (5500, 7000)$ MeV/c^2 and the lower mass sideband is $m_{K\pi\mu\mu} \in (5000, 5180)$ MeV/c^2. Although the upper mass sideband does have a slight overlap with the final analysis sample, it helps in the stability of the fit in the two regions. Two Gaussian distributions with common mean and common left side power law tail (double Crystal-Ball) are used for the fits to $B^0 \to J/\psi K^{*0}$ signal events and a double exponential is used for the fit to the background sidebands, example of these fits is shown in Figure 3.24. The fit quality is sufficient to obtain a realistic estimate of the background within the signal region.

The distribution of the significance as a function of the BDT cut is shown in Figure 3.25. There is a cluster of points between 0.1 and 0.2, all with effectively the same figure of merit. The tightest cut out of this cluster at 0.2 is chosen in order to keep background levels and therefore potential systematic effects under control. In addition, a tighter cut is preferable for the q^2 region [1,6] GeV/c^2 which has a lower signal over background ratio. The optimal working point for the BDT at 0.2 gives a signal efficiency of 85%, evaluated on $B^0 \to J/\psi K^{*0}$ candidates, and a background rejection rate of 97% in the $B^0 \to K^{*0}\mu^+\mu^-$ upper mass sideband. These performances are evaluated on the data sample for which the full remaining selection was applied.

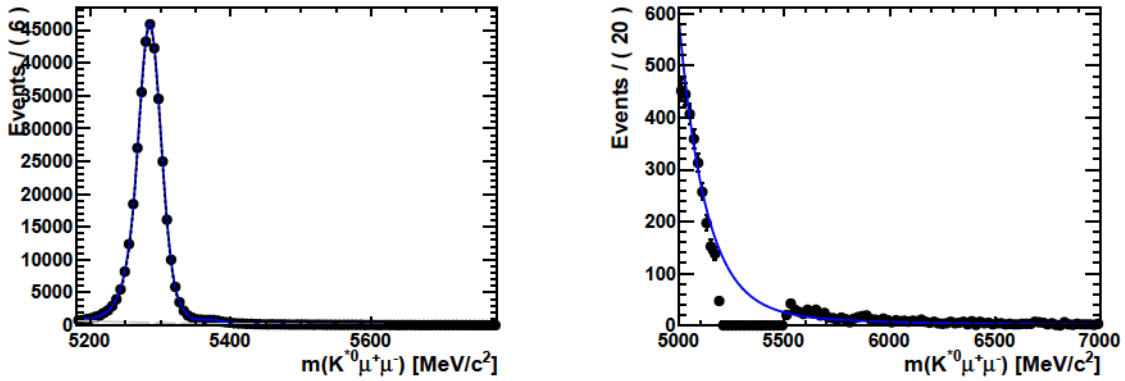


Figure 3.24: Example fits to $B^0 \to J/\psi K^{*0}$ (left) and $B^0 \to K^{*0}\mu^+\mu^-$ sidebands (right) used in the optimisation of the BDT cut as described in the text.

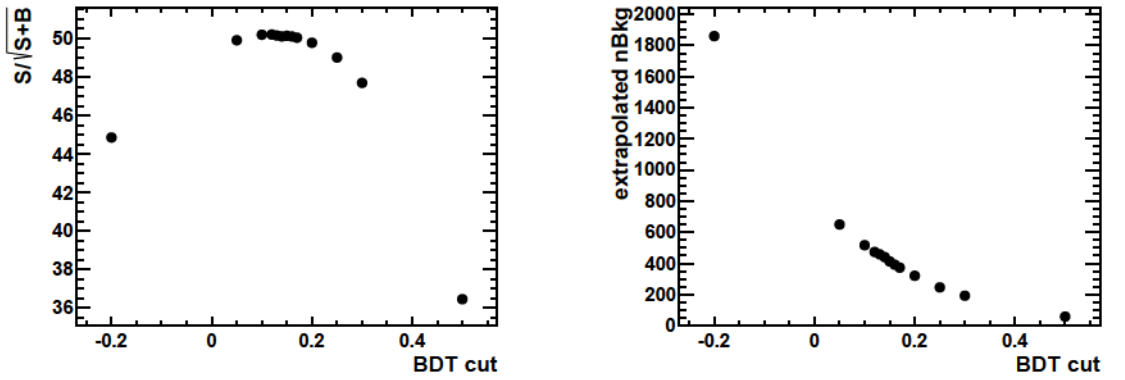


Figure 3.25: Significance (left) and background yield (right) as a function of the cut on the BDT variable.

3.6.6 Dependence of the BDT response on $m_{K\pi\mu\mu}$ and the angular distribution

The BDT is not directly given any information on the momenta of the daughters of the B^0 meson. Nevertheless several variables are correlated with these, such as PID or isolation variables, and can potentially give an indirect dependence of the BDT response on $m_{K\pi\mu\mu}$. This effect is studied by looking at the data events in the upper and lower mass sideband regions passing all the selection criteria apart from the BDT cut. Figure 3.26 shows the 2D plot of the BDT output versus the $m_{K\pi\mu\mu}$ across a wide $m_{K\pi\mu\mu}$ region including both lower and upper mass sidebands. A profile of the BDT response as a function of $m_{K\pi\mu\mu}$ is also shown. A small linear dependence of the response on $m_{K\pi\mu\mu}$ is observed with no evident structure above or below the signal mass region. Figure. 3.27 shows the 2D and profile plots for the upper mass sideband only. The efficiency of the BDT cut as a function of $m_{K\pi\mu\mu}$ is also shown. The correlation of $m_{K\pi\mu\mu}$ with the BDT cut in this region is 6%. Given the fact that we use a parametric model for the background, the dependence is smooth across the full mass region, and as the angular distribution of the background factorises with $m_{K\pi\mu\mu}$, this small correlation is acceptable.

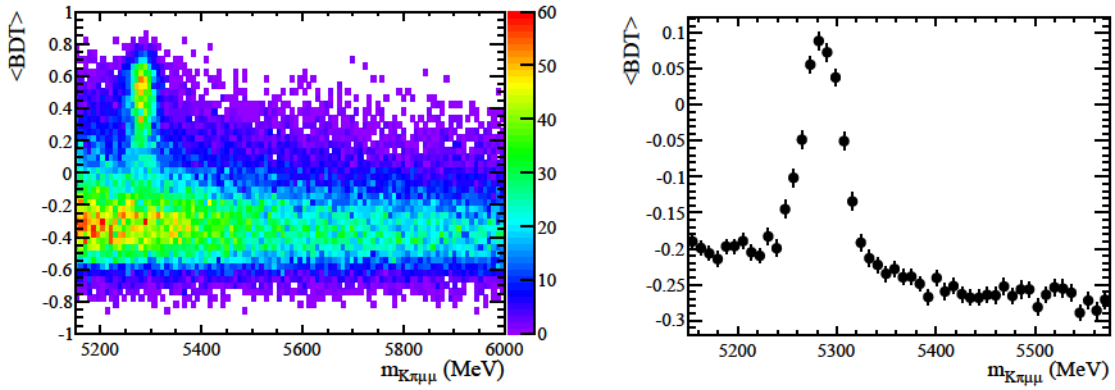


Figure 3.26: (left) The BDT response versus $m_{K\pi\mu\mu}$. (right) Profile of the mean (points) and RMS (error bars) of the BDT response in bins of $m_{K\pi\mu\mu}$.

Finally, a comparison of the efficiency of the 10 BDTs responses has been made for the three decay angles, $\cos\theta_l$, $\cos\theta_K$, ϕ, and q^2 and the $K^+\pi^-$ invariant mass. No differences is observed between the BDTs responses and a flat efficiency across the $K^+\pi^-$ invariant mass can be seen in Fig. 3.28.

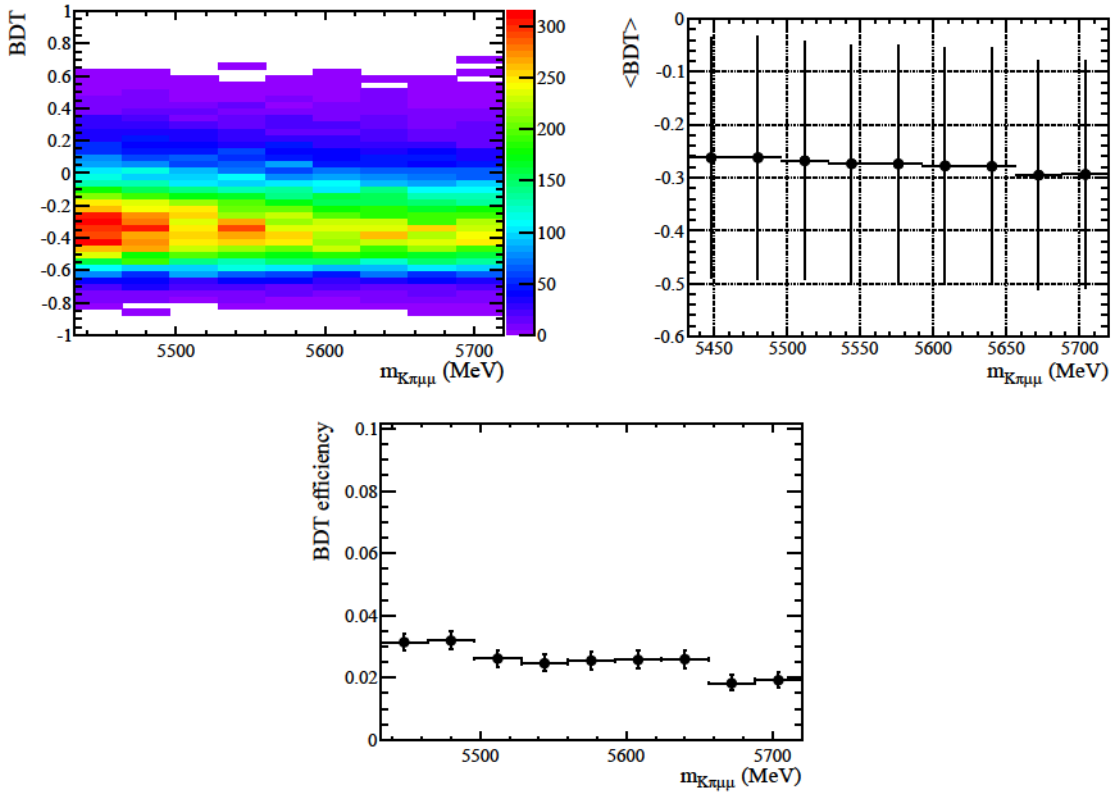


Figure 3.27: (left) The BDT response versus $m_{K\pi\mu\mu}$. (right) Profile of the mean (points) and RMS (error bars) of the BDT response in bins of $m_{K\pi\mu\mu}$. (bottom) Efficiency of the BDT cut with $m_{K\pi\mu\mu}$.

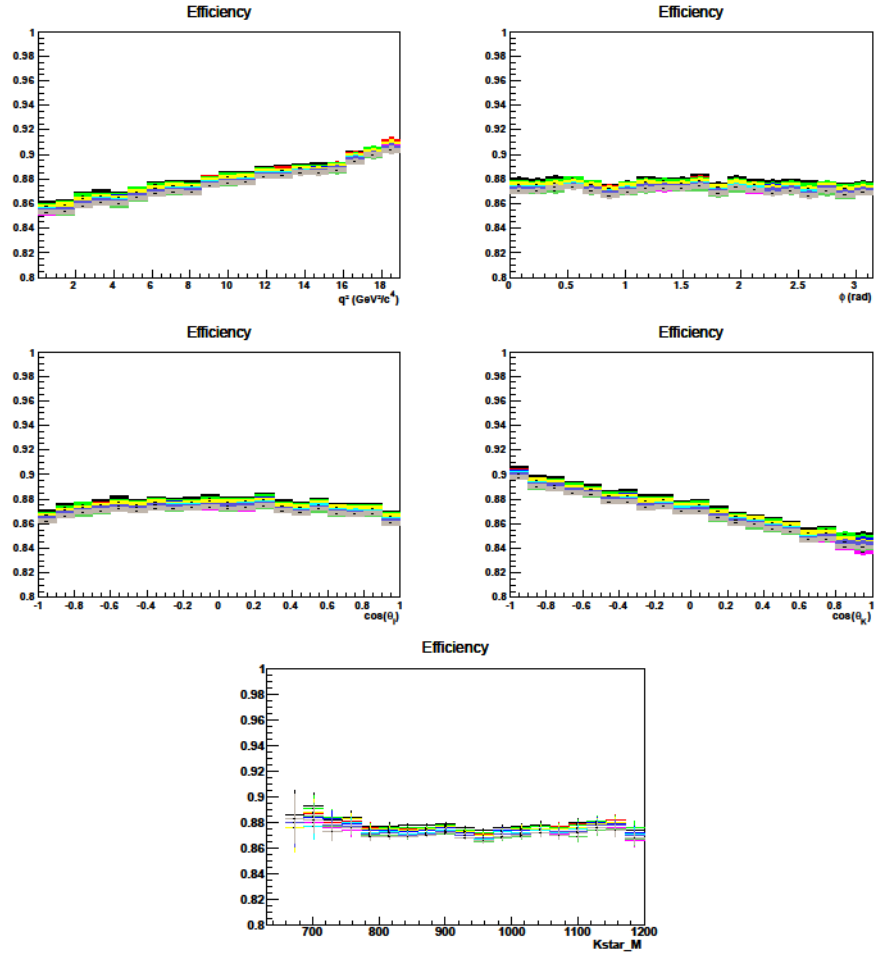


Figure 3.28: The variation of the efficiency of the 10 BDTs for the decay angles, q^2 and the $K^+\pi^-$ invariant mass.

REFERENCES

[1] T. Blake *et al.*, *Measurement of the full angular distribution of* $B^0 \to K^{*0}\mu^+\mu^-$ *decays using* $3\,\text{fb}^{-1}$ *of integrated luminosity*, LHCb-ANA-2013-097. Work in progress, at time of writing.

[2] S. Cunliffe, M. Patel, K. Petridis, and A. Shires, *Measurement of the differential branching fraction and S-wave fraction of* $B^0 \to K^+\pi^-\mu^+\mu^-$ *decays using* $3\,\text{fb}^{-1}$ *of integrated luminosity*, LHCb-ANA-2013-096. Work in progress, at time of writing.

[3] $B^0 \to K^{*0}\mu^+\mu^-$ WG, $B^0 \to K^{*0}\mu^+\mu^-$ *selection*, LHCb-INT-2013-058, 2013.

[4] T. Blake *et al.*, *Angular analysis of* $B^0 \to K^{*0}\mu^+\mu^-$ *at LHCb with 1* fb^{-1}, LHCb-ANA-2012-051.

[5] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[6] T. Blake, U. Egede, and A. Shires, *The effect of S-wave interference on the* $B^0 \to K^{*0}\ell^+\ell^-$ *angular observables*, JHEP **1303** (2013) 027, arXiv:1210.5279.

[7] C. Parkinson, M. Patel, and U. Egede, *The angular analysis of the* $B^0 \to K^{*0}\mu^+\mu^-$ *decay at LHCb*, PhD thesis, Imperial Coll., London, June, 2013, Presented 26 July 2013.

[8] S. Cunliffe, *The scalar component of* $B^0 \to K^+\pi^-\mu^+\mu^-$ *decays*, PhD thesis, Imperial College, London, September, 2015.

[9] T. Blake, U. Egede, P. Owen, and K. Petridis, *Measurement of the isospin asymmetry in* $B \to K^{(*)}\mu^+\mu^-$ *decays with 3* fb^{-1}, LHCb-ANA-2013-090.

[10] J. Lefrançois and M. H. Schune, *Measuring the photon polarization in b → s gamma using the B → K*e+e- decay channel*, Tech. Rep. LHCb-PUB-2009-008. CERN-LHCb-PUB-2009-008. LHCb-INT-2009-007, CERN, Geneva, Jun, 2009.

[11] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, European Physical Journal C **73** (2013) 2431, arXiv:1211.6759.

[12] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay* $B^0 \to K^{*0}\mu^+\mu^-$, JHEP **08** (2013) 131, arXiv:1304.6325.

[13] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[14] R. E. Schapire and Y. Freund, *A decision-theoretic generalization of on-line learning and an application to boosting*, Jour. Comp. and Syst. Sc. **55** (1997) 119.

[15] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. A **543** (2005) 577, arXiv:physics/0408124.

[16] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*, Ann. Statist. **29** (2001) 1189.

[17] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[18] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ Branching Fraction and Search for $B^0 \to \mu^+\mu^-$ Decays at the LHCb Experiment*, Phys. Rev. Lett. **111** (2013) 101805.

[19] G. Mancinelli and J. Serrano, *Study of Muon Isolation in the $B_s^0 \to \mu^+\mu^-$ Channel*, LHCb-INT-2010-011, 2010.

[20] A. Blum, A. Kalai, and J. Langford, *Beating the Hold-out: Bounds for K-fold and Progressive Cross-validation*, in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, COLT '99, (New York, NY, USA), pp. 203–208, ACM, 1999. doi: 10.1145/307400.307439.

Chapter 4

Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay

Sommaire

The candidates passing the selection requirements detailed in the previous chapter have been fit to measure the angular observables of interest for the $B^0 \to K^{*0}\mu^+\mu^-$ decay. An overview of the fit strategy is given in Section 4.1. The expression for the angular distribution is recalled in Section 4.2. The background angular distribution is discussed in Section 4.3, with a focus on the cross check performed to validate the analysis approach. In order to correctly describe the shape of the invariant mass of $B^0 \to K^{*0}\mu^+\mu^-$ signal events, some parameterisations, described in Section 4.5, have been studied. The mass model chosen enters directly the maximum likelihood fit of the angular distributions of the decay, helping to discriminate the residual background among the selected events. Finally, a possible contribution from a $K\pi$ S-wave is accounted for in Section 4.6.

4.1 The angular analysis strategy

The angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ presented here determines the angular observables S_i (or A_i) in bins of q^2 using an unbinned maximum likelihood fit of the reconstructed

B^0 mass, m, and the decay angles $\vec{\Omega} = (\cos\theta_l, \cos\theta_K, \phi)$. The observables and the different variables are defined in Sec. 1.3. This analysis has measured, for the first time, a complete set of CP-averaged observables using a global unbinned maximum likelihood fit. Thanks to this, the correlations between the observables can be computed, so that the measurements can be included in the global fits to theoretical models in a statistically correct way. This was not possible in the previous analysis [1, 2], since it was performed on a smaller dataset and so had to use some angular foldings of the data.

The analysis describes the signal and background components using probability density functions (PDFs) depending on the mass m and the angles $\vec{\Omega}$. The total PDF is given by:

$$\mathcal{P}_{\text{tot}} = f_{\text{sig}}\mathcal{P}_{\text{sig}}(\vec{\Omega}, m) + (1 - f_{\text{sig}})\mathcal{P}_{\text{bkg}}(\vec{\Omega}, m). \tag{4.1}$$

Both for signal and background, the mass and angular components are assumed to factorise:

$$\mathcal{P}_{\text{sig}}(\vec{\Omega}, m) = \mathcal{P}_{\text{sig}}(\vec{\Omega}) \times \mathcal{P}_{\text{sig}}(m) \tag{4.2}$$

$$\mathcal{P}_{\text{bkg}}(\vec{\Omega}, m) = \mathcal{P}_{\text{bkg}}(\vec{\Omega}) \times \mathcal{P}_{\text{bkg}}(m). \tag{4.3}$$

The reconstruction and selection cause a distortion of the angular distributions of the final state particles. This angular acceptance effects are taken into account in the signal PDF, described in Section. 4.4. To determine the angular observables, a negative logarithmic likelihood is built from the PDF:

$$-\log\mathcal{L} = -\sum_{\text{events } e} \log\mathcal{P}_{\text{tot}}(\vec{\Omega}_e, m_e|\vec{\lambda}_{\text{phys}}, \vec{\lambda}_{\text{nuisance}}) \tag{4.4}$$

and is minimised with respect to the physics parameters $\vec{\lambda}_{\text{phys}}$, which are the angular observables, and the nuisance parameters $\vec{\lambda}_{\text{nuisance}}$. The minimisation is performed using the MINUIT software package. Uncertainties on the parameters can be either determined using the second derivative matrix (HESSE) or the $-2\Delta\log\mathcal{L} = 1$ rule (MINOS), which allows asymmetric uncertainties.

It has to be noted that, in order to correctly determine the angular observables, the contribution from a $K\pi$ S-wave has to be accounted for. The strategy, described in detail in Section 1.3, makes use of a simultaneous fit of the mass and the angles distributions with the $m_{K\pi}$ distribution.

Throughout the analysis, the tree-level decay $B^0 \to J/\psi K^{*0}$ is used as a control-channel. It is an important cross-check for the description of the fit strategy.

The angular fit has measured the 8 CP-averaged observables: F_L, A_{FB} and $S_{3,4,5,7,8,9}$, described by Eq. 1.56, and the 7 CP-asymmetries $A_{3,4,5,6,7,8,9}$ defined in Eq. 1.57. The form-factor independent observables $P^{(\prime)}$ have also been measured.

These measurements have been performed in bins of q^2. With the increase of data available for the analysis, the traditional q^2 binning scheme used by the B-factories [3, 4] has been abandoned, and a new finer binning has been adopted. This has been decided on the basis of the available number of events in each region, the vetoes on the ϕ (Section 3.5.2) and on the charmonium resonances (Section 3.5.1) and recommendations from theoreticians. A 2 GeV$^2/c^4$ binning scheme, shown in the Table 4.1, is chosen for the maximum likelihood approach described here. Two larger q^2 bins, $q^2 \in [1,6]\,\text{GeV}^2/c^4$ and $q^2 \in [15,19]\,\text{GeV}^2/c^4$ have also been used, as some theory groups provide their predictions in these wider regions. It is worth mentioning here that another approach for the angular analysis has been performed within the LHCb collaboration, which is not described in this thesis: the method of moments. For this approach a 1 GeV$^2/c^4$ binning scheme, described in the Table 4.2, is used.

Bin	$q^2[\mathrm{GeV}^2]$
1	$[0.1, 0.98]$
2	$[1.1, 2.5]$
3	$[2.5, 4.0]$
4	$[4.0, 6.0]$
5	$[6.0, 8.0]$
6	$[11.0, 12.5]$
7	$[15.0, 17.0]$
8	$[17.0, 19.0]$

Table 4.1: The $2\,\mathrm{GeV}^2$ q^2 binning scheme.

Bin	$q^2[\mathrm{GeV}^2]$
1	$[0.1, 0.98]$
2	$[1.1, 2.0]$
3	$[2.0, 3.0]$
4	$[3.0, 4.0]$
5	$[4.0, 5.0]$
6	$[5.0, 6.0]$
7	$[6.0, 7.0]$
8	$[7.0, 8.0]$
9	$[11.0, 11.75]$
10	$[11.75, 12.5]$
11	$[15.0, 16.0]$
12	$[16.0, 17.0]$
13	$[17.0, 18.0]$
14	$[18.0, 19.0]$

Table 4.2: The $1\,\mathrm{GeV}^2$ q^2 binning scheme.

4.2 The signal angular distributions

The angular description of the signal component of the PDF is given by the differential decay rate in Eq. 1.52. The data are binned in q^2, thereby effectively averaging the observables over the width of the q^2 bins. The resulting three-differential decay rate is given by

$$\begin{aligned}\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\cos\theta_l\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi}\bigg|_{\mathrm{P}} = \tfrac{9}{32\pi}\big[&\tfrac{3}{4}(1-F_{\mathrm{L}})\sin^2\theta_K \\ &+ F_{\mathrm{L}}\cos^2\theta_K + \tfrac{1}{4}(1-F_{\mathrm{L}})\sin^2\theta_K\cos 2\theta_l \\ &- F_{\mathrm{L}}\cos^2\theta_K\cos 2\theta_l + S_3\sin^2\theta_K\sin^2\theta_l\cos 2\phi \\ &+ S_4\sin 2\theta_K\sin 2\theta_l\cos\phi + S_5\sin 2\theta_K\sin\theta_l\cos\phi \\ &+ \tfrac{4}{3}A_{\mathrm{FB}}\sin^2\theta_K\cos\theta_l + S_7\sin 2\theta_K\sin\theta_l\sin\phi \\ &+ S_8\sin 2\theta_K\sin 2\theta_l\sin\phi + S_9\sin^2\theta_K\sin^2\theta_l\sin 2\phi\big].\end{aligned} \tag{4.5}$$

As discussed in Sec. 1.3.3.1, the inclusion of an S-wave contribution, where the $K^+\pi^-$ system is in a spin 0 configuration, leads to additional angular terms. The PDF needs to be changed as described by the Eq. 1.61

4.3 Study of the background angular distributions

In this analysis, the background angular component in the maximum likelihood fit is modeled using Chebyshev polynomials. The angular parametrisation of the background is assumed to factorise in each of its components. Using Chebyshev polynomials T_i of second order and lower, the PDF for the angular parametrisation of the background is given by

$$\mathcal{P}_{\mathrm{bkg}}(\cos\theta_l,\cos\theta_K,\phi) = \left[\sum_{i=0}^{2} c_i T_i(\cos\theta_l)\right]\times\left[\sum_{j=0}^{2} c_j T_j(\cos\theta_K)\right]\times\left[\sum_{k=0}^{2} c_k T_k(\phi)\right] \tag{4.6}$$

The coefficients are determined from the fit in the whole region. It assumes that events in the upper mass sideband have the same angular distribution as those in the signal region. To test and to validate this assumption, the ABCD method [5] is used as a crosscheck. A comparison between the background angular distribution determine from the PDF 4.6 and the ABCD method is made in this section.

4.3.1 Parametrisation cross check using the ABCD method

The ABCD method exploits different regions of the BDT versus B^0 mass plane. Namely:

A: BDT > 0.2 and $m(K^+\pi^-\mu^+\mu^-) < 5350\,\text{MeV}/c^2$

B: BDT > 0.2 and $m(K^+\pi^-\mu^+\mu^-) > 5350\,\text{MeV}/c^2$

C: BDT < -0.4 and $m(K^+\pi^-\mu^+\mu^-) < 5350\,\text{MeV}/c^2$

D: BDT < -0.4 and $m(K^+\pi^-\mu^+\mu^-) > 5350\,\text{MeV}/c^2$

The goal of the ABCD method is to infer the distribution of the background in the region A, referred as "the the signal region", using the distributions from three control regions (B - D), where the region B to the commonly called "upper mass sideband region".

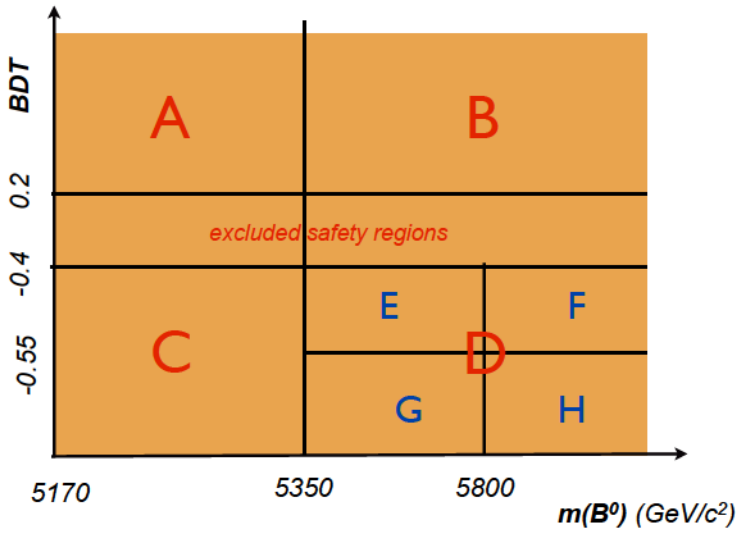


Figure 4.1: 2D plane defining the different regions of the ABCD method.

The distribution in the region A is obtained from the distribution of the regions B, C and D by:

$$A = \frac{B \times C}{D}, \tag{4.7}$$

were here A, B, C and D represent the angular distribution for each region.

The advantage of the ABCD approach is that there is no assumption that the angular distribution of the background is the same in the sideband and in the signal mass window. The ABCD method can then be used to test this assumption.

The method relies on two hypothesis:

1. the regions B, C, D contain only background events;
2. there is no correlation between the BDT variable and the reconstructed mass of the candidate.

To ensure that the first hypothesis is true, events in a safety zone are excluded from the regions B and D, this zone being defined by an interval in the MVA response from -0.4 to 0.2. The leakage of signal events into region B is then at most one event in every q^2 bin.

Concerning the second hypothesis, unfortunately, when the BDT requirement is relaxed, there can be a correlation between the BDT response and the mass of the candidate. To take into account this correlation, a correction to the angular distribution in the A region is applied. The correction factor is computed as the following:

$$R = \frac{E \times H}{F \times G} \tag{4.8}$$

where E,F,G and H are the number of events of subsets of the region D, defined as:

E: $-0.55 < \texttt{BDT} < -0.4$ and $5350 < m(K^+\pi^-\mu^+\mu^-) < 5800\,\text{MeV}/c^2$;

F: $-0.55 < \texttt{BDT} < -0.4$ and $m(K^+\pi^-\mu^+\mu^-) > 5800\,\text{MeV}/c^2$;

G: $\texttt{BDT} < -0.55$ and $5350 < m(K^+\pi^-\mu^+\mu^-) < 5800\,\text{MeV}/c^2$;

H: and $\texttt{BDT} < -0.55$ and $m(K^+\pi^-\mu^+\mu^-) > 5800\,\text{MeV}/c^2$.

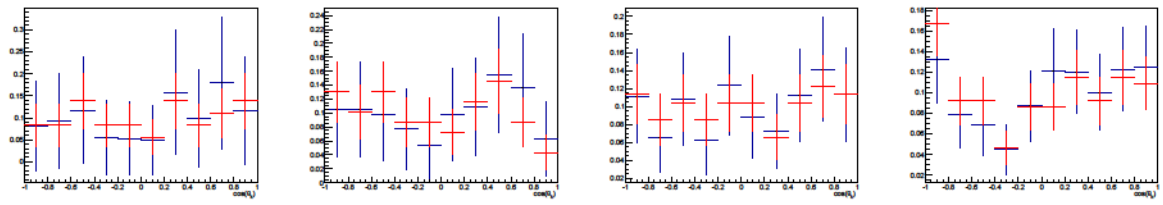
The correction factors for the different q^2 bins are given in Table 4.3. It has been checked that this correction factor has a negligible angular dependence, see Appendix. C.1.

The comparison of the background angular distributions in the A region, obtained from the ABCD method, to that from the upper mass sideband is shown for the three angles in Figs. 4.2, 4.3, 4.4, and 4.5. The same comparison has also been done for the smaller binning scheme and is shown in Appendix. C.2. The distributions are in excellent agreement, and this study gives confidence on the description of the background angular distributions used in the analysis.

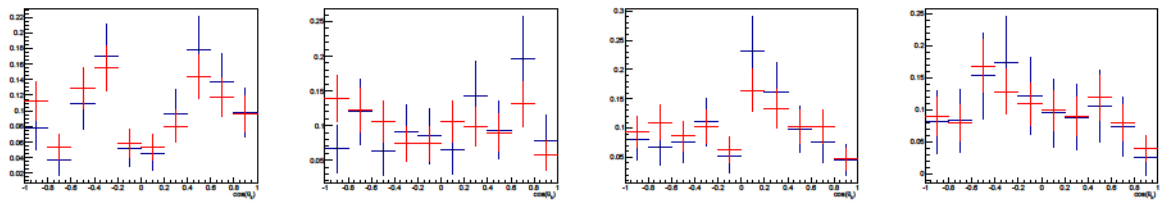
$q^2[\mathrm{GeV}^2]$	Correction factor R
[0.1, 0.98]	1.304
[1.1, 2.0]	1.584
[2.0, 3.0]	1.259
[3.0, 4.0]	1.366
[4.0, 5.0]	1.281
[5.0, 6.0]	1.276
[6.0, 7.0]	1.324
[7.0, 8.0]	1.282
[11.0, 11.75]	1.435
[11.75, 12.5]	1.007
[15.0, 16.0]	1.438
[16.0, 17.0]	0.994
[17.0, 18.0]	1.057
[18.0, 19.0]	0.902

$q^2[\mathrm{GeV}^2]$	Correction factor R
[0.1, 0.98]	1.034
[1.1, 2.5]	1.517
[2.5, 4.0]	1.301
[4.0, 6.0]	1.282
[6.0, 8.0]	1.304
[11.0, 12.5]	1.206
[15.0, 17.0]	1.176
[17.0, 19.0]	0.968

Table 4.3: Correction factor R within the two q^2 binning schemes: (left) $1\,\mathrm{GeV}^2/c^4$ scheme and (right) $2\,\mathrm{GeV}^2/c^4$ scheme.

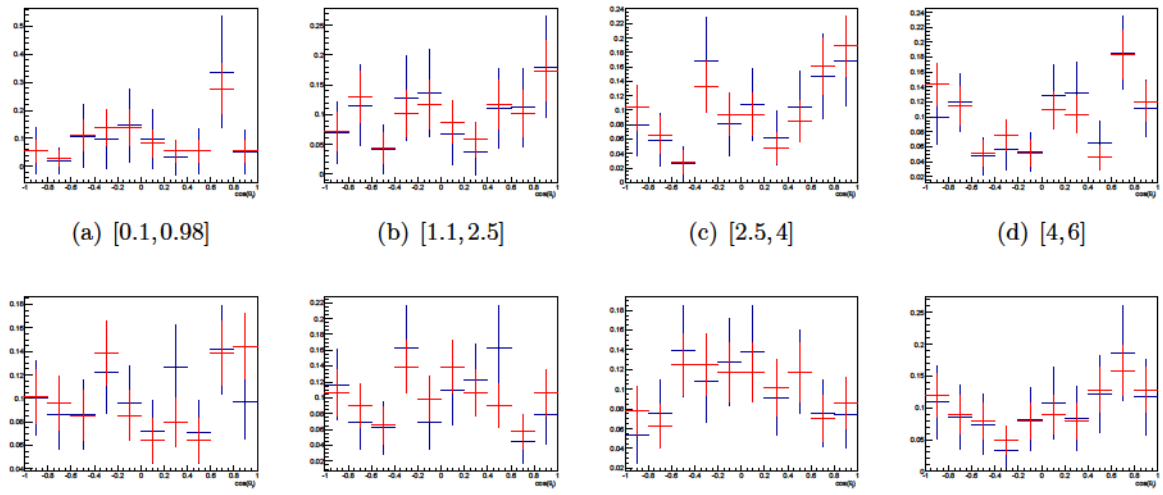


(a) [0.1, 0.98] (b) [1.1, 2.5] (c) [2.5, 4] (d) [4, 6]



(e) [6, 8] (f) [11, 12.5] (g) [15, 17] (h) [17, 19]

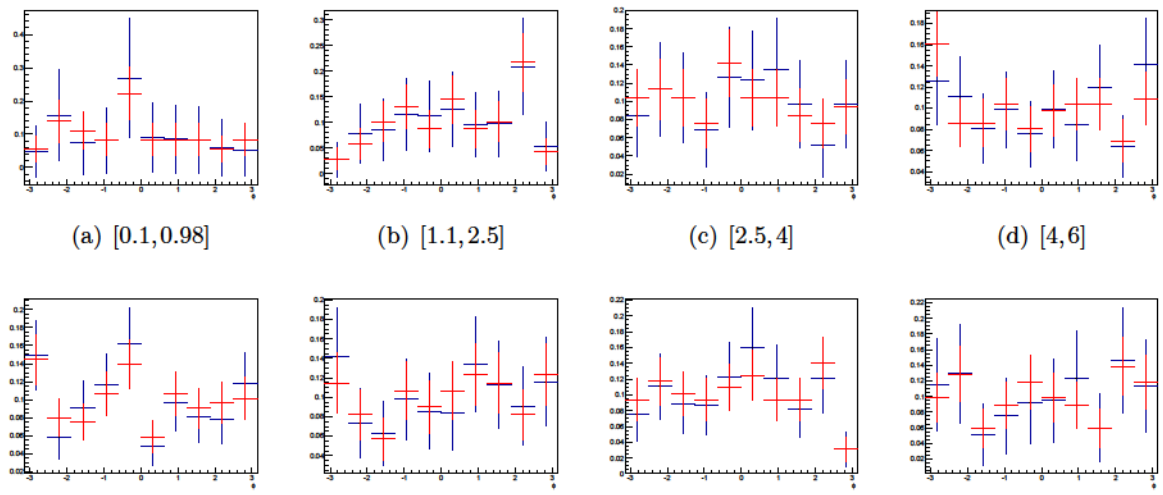
Figure 4.2: Comparison of the $\cos(\theta_k)$ background distributions obtained from the ABCD method (blue) and simply from the upper mass sideband (red) in the 2 GeV^2/c^4 binning scheme.



(a) [0.1,0.98] (b) [1.1,2.5] (c) [2.5,4] (d) [4,6]

(e) [6,8] (f) [11,12.5] (g) [15,17] (h) [17,19]

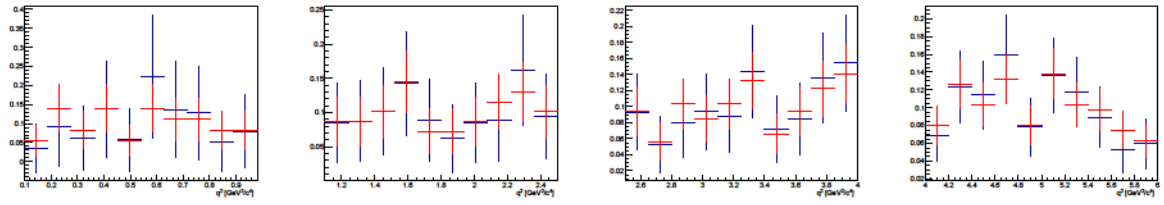
Figure 4.3: Comparison of the $\cos(\theta_l)$ background distributions obtained from the ABCD method (blue) and simply from the upper mass sideband (red) in the 2 GeV^2/c^4 binning scheme.



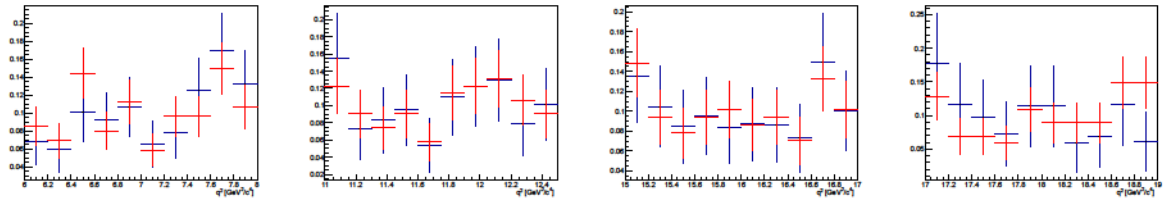
(a) [0.1,0.98] (b) [1.1,2.5] (c) [2.5,4] (d) [4,6]

(e) [6,8] (f) [11,12.5] (g) [15,17] (h) [17,19]

Figure 4.4: Comparison of the ϕ background distributions obtained from the ABCD method (blue) and simply from the upper mass sideband (red) in the 2 GeV^2/c^4 binning scheme.

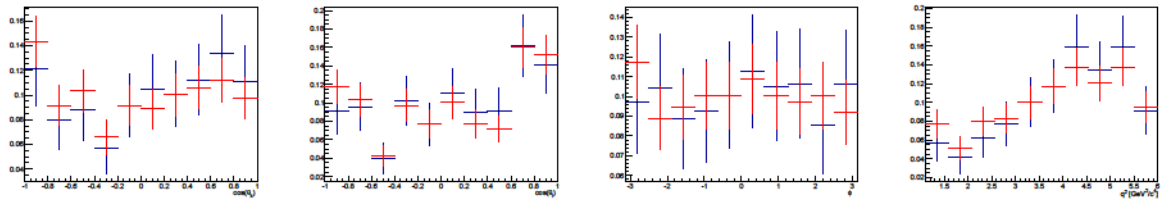


(a) [0.1, 0.98] (b) [1.1, 2.5] (c) [2.5, 4] (d) [4, 6]



(e) [6, 8] (f) [11, 12.5] (g) [15, 17] (h) [17, 19]

Figure 4.5: Comparison of the q^2 distribution for the ABCD method (blue) and simply from the upper mass sideband (red) in the 2 GeV$^2/c^4$ binning scheme.



(a) $\cos(\theta_k)$ (b) $\cos(\theta_l)$ (c) ϕ (d) q^2

Figure 4.6: Comparison of the $\cos(\theta_k)$ 4.6(a) ,$\cos(\theta_l)$ 4.6(b), ϕ 4.6(c) and q^2 4.6(d) background distributions obtained from the ABCD method (blue) and simply from the upper mass sideband (red) in the large bin 1-6 GeV$^2/c^4$.

4.4 Impact of the selection acceptance on the angular distributions

The reconstruction, trigger and the selection cuts distort the angular distributions of the signal and need to be accounted for in the fit when measuring the angular observables. For example, the cut on the p_T of the muons leads to a drop in the $\cos(\theta_\ell)$ distribution for values around ± 1 in the first q^2 bin; another example is the observed asymmetric distribution of the $\cos(\theta_k)$ variable, related to the p_T cuts having a different impact on the kaon and the pion, due to their different mass.

The shape of the acceptance of the full selection for the three decay angles and q^2 is shown by the distribution in Figure 4.7, obtained from a phase space Monte Carlo simulation. These acceptances can be parameterised using multidimensional polynomials:

$$\varepsilon(\cos\theta_\ell, \cos\theta_K, \phi, q^2) = \sum_{k,l,m,n} c_{k,l,m,n} P(\cos\theta_\ell, k) P(\cos\theta_K, l) P(\phi, m) P(q^2, n). \tag{4.9}$$

where the terms $P(x,i)$ are Legendre polynomials of order i and the $c_{k,l,m,n}$ coefficients are determined using a moment analysis of $B^0 \to K^{*0}\mu^+\mu^-$ phase space simulated events. For q^2 a seventh order polynomial is used, for $\cos\theta_\ell$ a polynomial of fifth order, while for both ϕ and $\cos\theta_K$ polynomials of sixth order. The solid lines in Figure 4.7 give the projections of the four-dimensional polynomial parametrisation of the acceptance. The determination of the polynomial coefficients and the resulting angular description is discussed in Sec. **??**.

The acceptance can be included in the fit in two ways: either by performing a weighted fit, in which the events are weighted by $1/\epsilon$; or by including the effect in the signal PDF.

In the first option, the distributions are effectively unfolded, therefore the original signal PDF without acceptance can be used. It should be noted that the background component will be weighted in the same way. The per-event weight is included in the likelihood as follows

$$\begin{aligned}\mathcal{L} &= -\sum_{\text{event } e} w_e \times \log \mathcal{P}(\vec{\Omega}_e, m_e) \\ &= -\sum_{\text{event } e} \frac{1}{\epsilon(q_e^2, \vec{\Omega}_e)} \times \log \mathcal{P}(\vec{\Omega}_e, m_e).\end{aligned}$$

Special care needs to be taken for the estimation of the parameter uncertainties, since weighted fits in general are not guaranteed correct coverage. However, an approximate methods exists. The corrected covariance matrix V' for the weighted fit can be calculated according to

$$V' = VC^{-1}V,$$

where V is the covariance matrix calculated with the weights w_e and C the covariance matrix calculated using the squared weights w_e^2 [6]. The unfolding using acceptance weights is the preferred approach for the large q^2 bins $1.1 < q^2 < 6$ GeV$^2/c^4$ and $15 < q^2 < 19$ GeV$^2/c^4$, since the method can account for possible variation of the acceptance with q^2. Furthermore, the expected signal yield in these bins is sufficiently large to reduce possible fluctuations from the weighting procedure.

The second option requires to include the efficiency in the signal PDF. The main difficulty with this approach is the correct determination of the normalization of the signal component

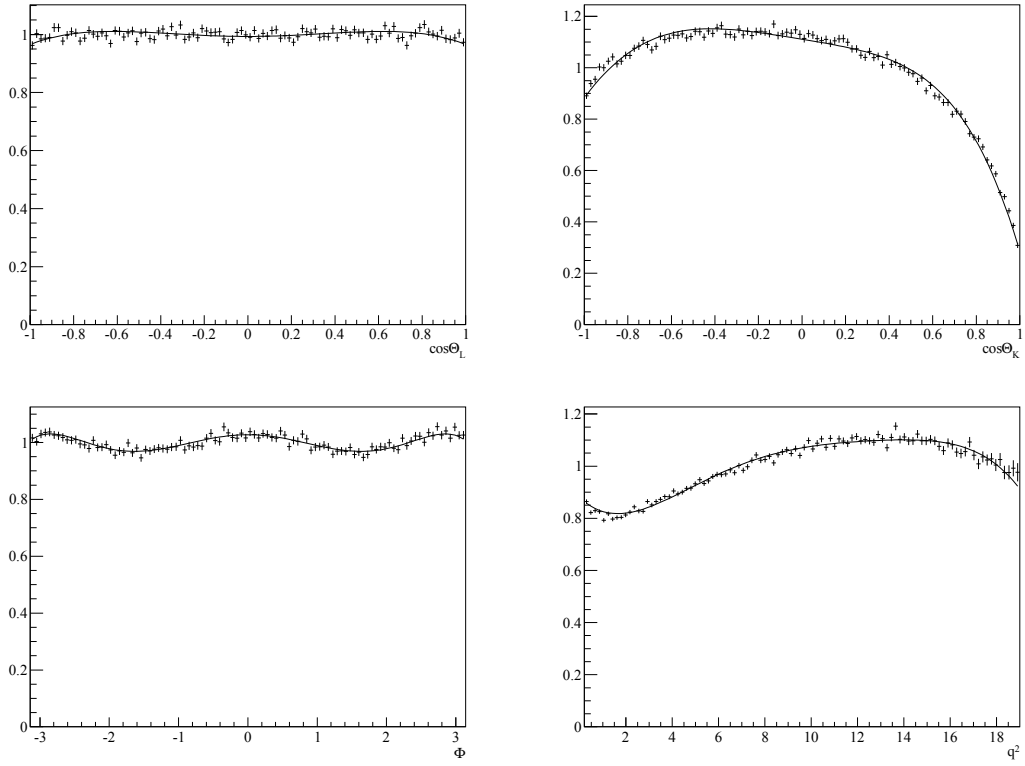


Figure 4.7: The efficiency of the full selection on the decay angles and q^2 as determined using fully simulated $B^0 \to K^{*0}\mu^+\mu^-$ events. The solid line gives a four-dimensional parametrisation of the efficiency using Chebyshev polynomials. The total scale is arbitrary.

which will be affected by the acceptance[1]. This normalization $\mathcal{N}_{\text{sig}}$ willl be given by

$$\begin{aligned}\mathcal{N}_{\text{sig}} &= \int \epsilon(q^2,\vec{\Omega})\mathcal{P}_{\text{sig}}(\vec{\Omega})\mathrm{d}\vec{\Omega} \\ &= \int \epsilon(q^2,\vec{\Omega})\frac{9}{32\pi}\sum_i S_i f_i(\vec{\Omega})\mathrm{d}\vec{\Omega} \\ &= \frac{9}{32\pi}\sum_i S_i \xi_i(q^2),\end{aligned} \tag{4.10}$$

with $\xi_i = \int \epsilon(q^2,\vec{\Omega}) f_i(\vec{\Omega})\mathrm{d}\vec{\Omega}$, and where $\vec{\Omega} = (\cos\theta_l, \cos\theta_k, \phi)$ and the angular terms $f_i(\vec{\Omega})$ are defined by the Eqs. 4.5 and 1.65. This is the preferred approach for the $2\,\text{GeV}^2/c^4$ q^2 bins, where the acceptance does not change too much inside each bin.

[1]Note that the factor $\epsilon(q^2,\vec{\Omega})$ in the numerator can be omitted when determining $-\log\mathcal{L}$.

4.5 Parameterisation of the signal invariant mass distribution

Even after the selection procedure described in the previous chapter, some background events are still included in the dataset on which the angular analysis is performed. In order to better disentangle the contribution of the signal and from background in the angular distributions, we multiply the angular PDF by a $K\pi\mu\mu$ invariant mass PDF, as described in Section 4.1. Since the final background discrimination rely on this term, it is crucial to provide a precise parameterisation of the mass. The strategy for modeling the $K\pi\mu\mu$ invariant mass shape of the $\mathrm{B}_d \to K^{*0}\mu\mu$ candidates is to exploit the large sample of $\mathrm{B}^0_d \to J/\psi K^{*0}$ in data. Monte Carlo (MC) simulated $\mathrm{B}_d \to K^{*0}\mu\mu$ and $\mathrm{B}_d \to J/\psi K^{*0}$ events are then used to study (and correct for) possible q^2 dependence of the mass shape parameters.

The following Sections 4.5.2 and 4.5.3 will present the different mass models that have been tested. The model currently used in the angular analysis will be presented in the Section 4.5.4. The event yields for the two q^2 binning schemes will be presented in Section 4.5.5.

4.5.1 The Crystal Ball distribution

All models tested for the $\mathrm{K}\pi\mu\mu$ invariant mass were based on a Crystal Ball (CB) function [7], named after the Crystal Ball Collaboration. This distribution describes an invariant mass spectrum with a peak and a radiative tail. The analytical expression of the Crystal Ball is given by:

$$\mathcal{P}_{\mathrm{CB}}(m|\mu,\sigma,\alpha,n) = N \times \begin{cases} e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2} & \frac{m-\mu}{\sigma} > -\alpha \\ \frac{a}{\left(b-\left(\frac{m-\mu}{\sigma}\right)\right)^n} & \frac{m-\mu}{\sigma} < -\alpha \end{cases}, \tag{4.11}$$

where:

$$\begin{aligned} a &= \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{1}{2}\alpha^2} \\ b &= \frac{n}{|\alpha|} - |\alpha| \ . \\ N &= \frac{1}{\sigma\left(\frac{n}{|\alpha|}\frac{1}{n-1}e^{-\frac{|\alpha|^2}{2}} + \sqrt{\frac{\pi}{2}}\left(1+\mathrm{erf}\left(\frac{|\alpha|}{2}\right)\right)\right)} \end{aligned} \tag{4.12}$$

In practice, the Crystal Ball is a Gaussian distribution above $m = -\alpha\sigma$, and a power law function below.

4.5.2 First parametrisation: two Crystal Ball functions with opposite tails

We first tested a sum of two Crystal Ball functions (DCB) with common mean (μ) but different widths and tails parameters, since it tends to correctly reproduce the $\mathrm{K}\pi\mu\mu$ distribution in simulated events(see Fig. 4.8). Explicitly, the reconstructed B_d mass was parameterised as:

$$\mathcal{P}_{\mathrm{sig}}(m|\vec{\lambda}) = f_{\mathrm{core}}\mathcal{P}_{\mathrm{CB}}(m|\mu,\sigma_1,\alpha_1,n_1) + (1-f_{\mathrm{core}})\mathcal{P}_{\mathrm{CB}}(m|\mu,\sigma_2,\alpha_2,n_2). \tag{4.13}$$

where f_{core} is the relative fraction of candidates falling in the first Crystal Ball function.

In order to study the q^2 dependence of the parameters of the DCB, we first applied individual fits in narrow q^2 bins on $B^0 \to K^{*0}\mu^+\mu^-$ MC events, and we plot the results for each parameter as a function of the q^2 of the bin. Then, we fit the results for each parameter with a linear

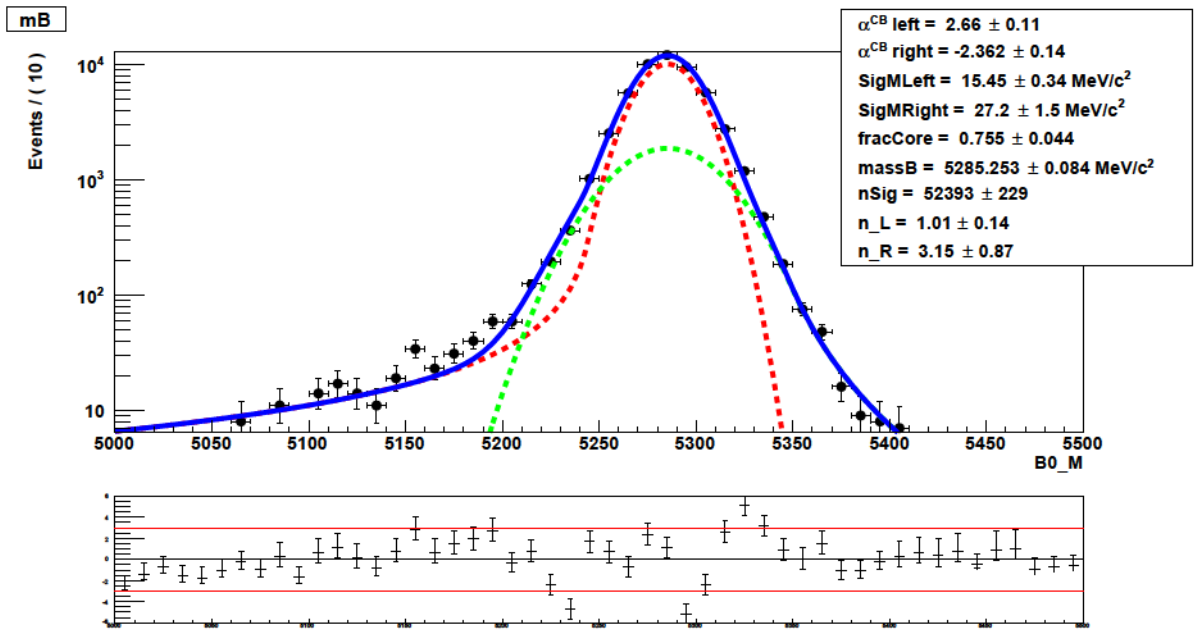


Figure 4.8: Fit with a DCB function to K$\pi\mu\mu$ mass distribution of $B_d \to J/\psi K^{*0}$ simulated events.

function:

$$X = A_X \times q^2 + B_X, \tag{4.14}$$

Here X represents a parameter of the Double Crystal Ball, while A_X and B_X are the coefficients to be determined that are describing the q^2 dependence of the parameter X.

As an alternative approach, we also performed a simultaneous fit over the different q^2 bins. The comparison of the q^2 dependencies obtained for the individual-bin fits and the simultaneous fit, for each of the parameters, are shown in Figure 4.9. The results of this comparison can be summarized as follows:

- there is a good agreement in the q^2-dependency parameters found by the individual and simultaneous fit;
- the simultaneous fit has smaller uncertainty, due to the fact that it exploits the full data sample, leading to a more precise determination of the parameters in each bin;
- the yields are stable and in agreement among the two methods.

4.5.3 Second parameterization: a modified Crystal Ball function

In order to reduce the number of parameters needed to model the $K\pi\mu\mu$ invariant mass, another parametric function has been tested. This function was inspired from the Double Crystal Ball, and has a gaussian core and a right and left power-law functions to describe the tails :

$$\mathcal{P}(m|\mu,\sigma,\alpha_L,n_L,\alpha_R,n_R) = \begin{cases} e^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2} & -|\alpha_L| <= \frac{m-\mu}{\sigma} <= |\alpha_R| \\ \frac{\left(\frac{n_{R,L}}{|\alpha_{R,L}|}\right)^{n_{R,L}} e^{-\frac{1}{2}\alpha_{R,L}^2}}{\left(\frac{n_{R,L}}{|\alpha_{R,L}|} - |\alpha_{R,L}| \pm \left(\frac{m-\mu}{\sigma}\right)\right)^{n_{R,L}}} & \frac{m-\mu}{\sigma} < -|\alpha_L| \text{ or } \frac{m-\mu}{\sigma} > |\alpha_R| \end{cases}, \tag{4.15}$$

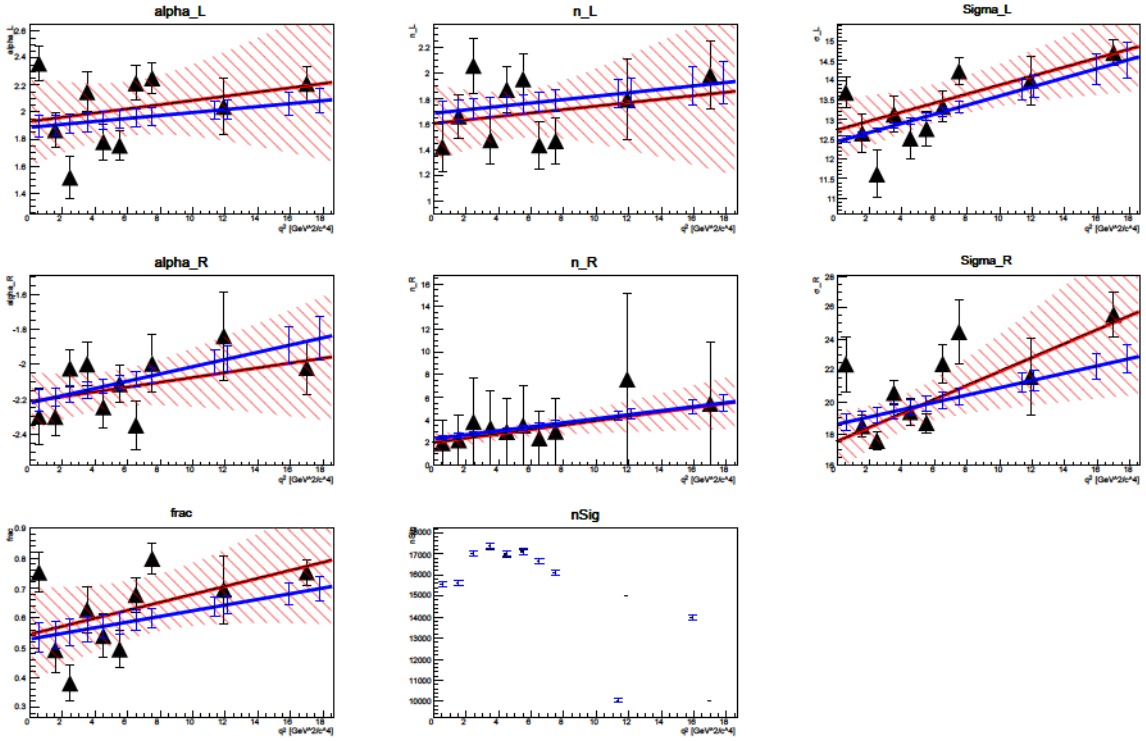


Figure 4.9: Double Crystal Ball parameters as a function of q^2. The measurements in black are the values obtained in the fits in each individual q^2 bin on simulated events. The lines and error bands in red are the result of the fit of these individual measurements. It has to be compared to the blue line giving the result of the simultaneous fit.

Following the methodology described in the previous Section 4.5.2, individual bin fits and a simultaneous fit have been performed to extract the 6 parameters of our model. The results of the simultaneous fits, represented for the data in only one of the bins, can be seen in Figure B.1 as example (the result for all the other bins are in the Appendix B). The q^2-dependency of the parameters is shown in Figure 4.11 for both individual bin and simultaneous fits.

Due to the small number of events in some of the bins, some deviation appears between the value of an individual fit in these bins and the one obtained by the simultaneous fit. These deviations are due to the correlation between the α and n parameter and this has been confirmed by fixing the α parameter in the problematic bins and comparing the new value for the n parameter from the individual fit with the one from the simultaneous fit, which now are in agreement (see blue dots in Figure 4.11).

Comparing the results from fits on the individual bins and the simultaneous fit, we end up with the same conclusion as for the previous parametrisation: the results of both fits are in agreement, but the simultaneous fit has a smaller error and provides a better control of the correlations between the parameters, leading to a more stable fit.

Since the MC does not take into account the detector resolution, we cannot directly apply the q^2 dependencies previously found in the fit to data. To overcome this issue we evaluate a scale factor for each parameter using the control decay $B^0 \to J/\psi K^{*0}$ and as a cross check the second charmonium resonance: $\psi(2S)$. In these fits of the regions containing charmonia in data, the combinatorial background is described by an exponential distribution and a second signal

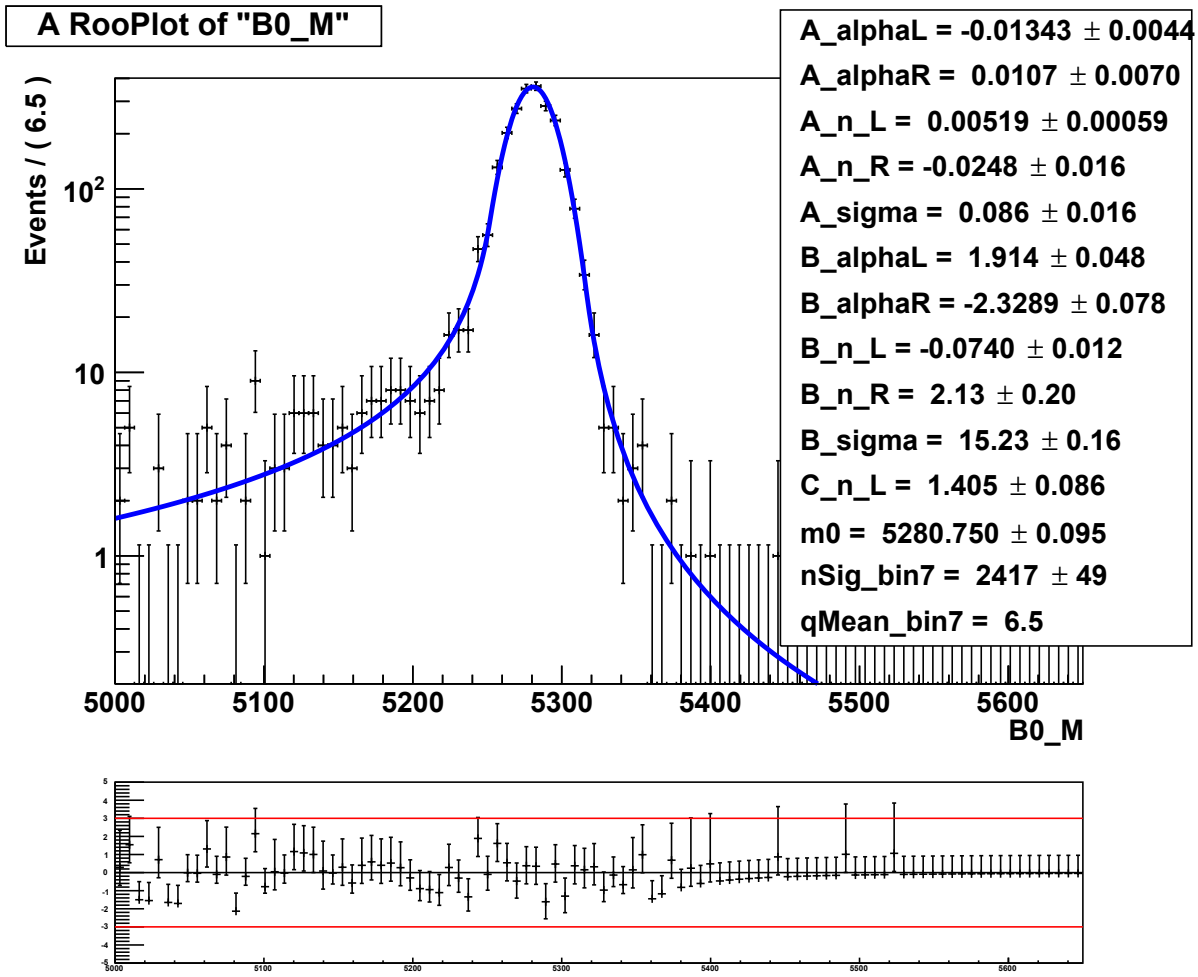


Figure 4.10: Results of simultaneous fit, in one of the q^2 bin, on the $K\pi\mu\mu$ mass distribution fit of $\mathrm{B}_d \to K^{*0}\mu\mu$ simulated events.

component is included for the $B_s^0 \to J/\psi\overline{K^{*0}}$ or $B_s^0 \to \psi(2S)\overline{K^{*0}}$ decays, expressed with the same signal parametrisation but with a shift on the mean (μ) by $\Delta m = m(B_s) - m(B^0)$. The scale factor obtained in the J/ψ region gives a good result on the $\psi(2S)$ region, validating the scaling approach (see Figure 4.12).

This parameterisation seems to work well. Nevertheless, some concerns are given by the fact that the right tail could be overestimated in presence of the $\mathrm{B}_s \to J/\psi\overline{K^{*0}}$ decay. For this reason, a more conservative approach has been chosen, described in the following section. In spite of this, the present study has clearly shown the need for scaling at least the widths of the CB as a function of q^2.

4.5.4 Third parametrisation: a Double Crystal Ball with widths depending on q^2

For the angular analysis, we have finally chosen to describe the $\mathrm{K}\pi\mu\mu$ invariant mass, m, by the sum of two Crystal Ball functions with a common mean (μ) and tail parameters (α and n)

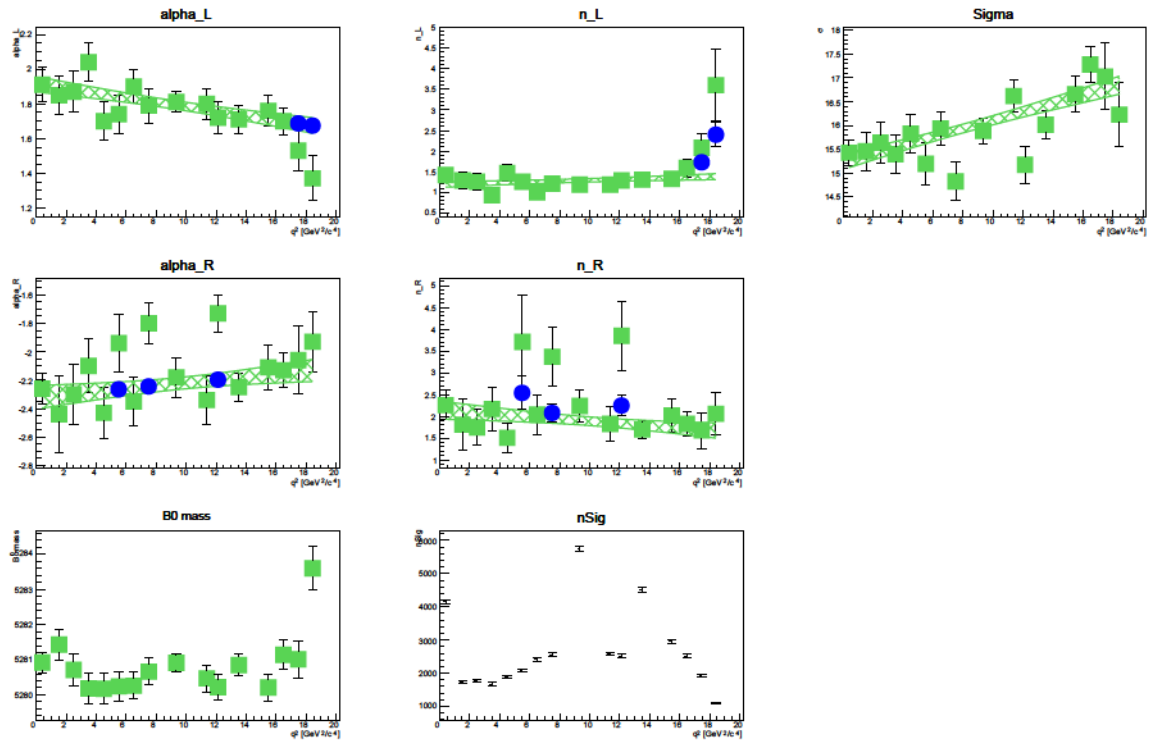


Figure 4.11: Parameters of the modified Crystal Ball as a function of q^2. The green squares are the value for the fits in each individual q^2 bin on simulated events. The green band is the result of the simultaneous fit. The new value of the n parameter in an individual fit when α is fixed to the result of the simultaneous fit is superposed with blue dots.

but different widths. Explicitly, the reconstructed B_d mass is parameterised as

$$\mathcal{P}_{\text{sig}}(m|\vec{\lambda}) = f_{\text{core}}\mathcal{P}_{\text{CB}}(m|\mu,\sigma_1,\alpha,n) + (1-f_{\text{core}})\mathcal{P}_{\text{CB}}(m|\mu,\sigma_2,\alpha,n). \tag{4.16}$$

The parameters to describe the $B^0_d \to K^{*0}\mu\mu$ signal mass shape, are determined from a fit to the control decay $B^0_d \to J/\psi K^{*0}$, which is shown in Fig. 4.13, where the invariant distribution of the background is described by an exponential distribution and where the $B^0_s \to J/\psi\overline{K^{*0}}$ decay is expressed with the same signal parametrisation and a shift on the mean (μ) by $\Delta m = m(B^0_s) - m(B^0_d)$. The resulting mass parameters for the angular analysis of the signal decay $B^0_d \to K^{*0}\mu\mu$ are given in Tab. 4.4.

To account for possible changes of the signal mass shape with q^2, we include a single scaling factor s_σ for every q^2 bin. This is applied to both the widths σ_1 and σ_2. The scaling factor is determined from a fit to MC simulated $B_d \to K^{*0}\mu\mu$ signal events. Its behaviour is shown in Fig. 4.14 while the numerical values are given in the Tab. 4.5.

4.5.5 Event yields

The $K\pi\mu\mu$ invariant mass distribution of $B^0_d \to K^{*0}\mu\mu$ candidates for the different q^2 bins are shown in Fig. 4.15, Fig. 4.16, together with the results of the fits. Table. 4.6 lists the signal and background yields in each q^2 bin for the two binning schemes. In total, 2390 signal candidates are seen within the range $0.1 < q^2 < 19$ GeV$^2/c^4$.

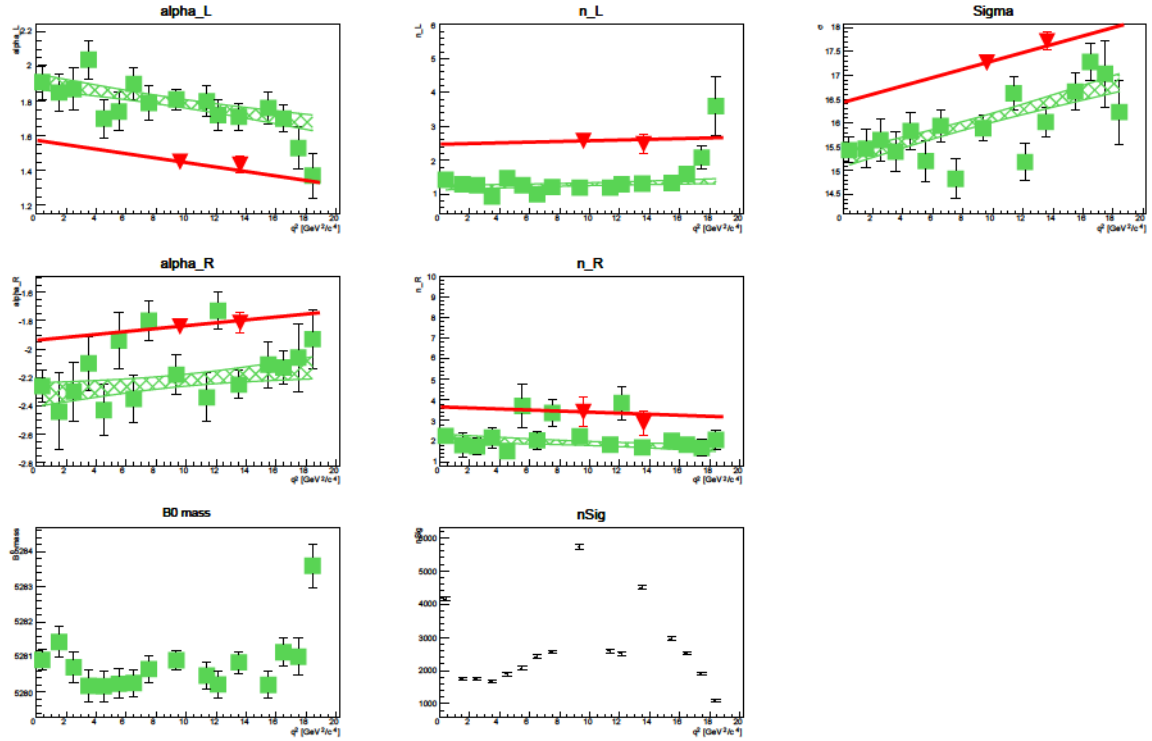


Figure 4.12: Parameters of the modified Crystal Ball as a function of q^2. The green squares are the value for the fits in each individual q^2 bin on simulated events. The green band is the result of the simultaneous fit. The results of the individual fits in the $J/\psi(1S)$ and $\psi(2S)$ regions, used for the data-MC rescaling, are shown in red.

Parameter	Value
α_{CB}	1.533 ± 0.033
n	4.23 ± 0.6
σ_1	15.36 ± 0.19
σ_2	25.85 ± 0.82
f_{core}	0.704 ± 0.031
B_d mass (μ)	5284.339 ± 0.043
Δm	87.21 ± 0.83
exp. slope	-0.006319 ± 0.0001
N_{sig}	343763 ± 822
N_{B_s}	4199 ± 162
N_{bkg}	26877 ± 649

Table 4.4: Value of the fit to the control channel $\text{B}_d \to J/\psi K^{*0}$.

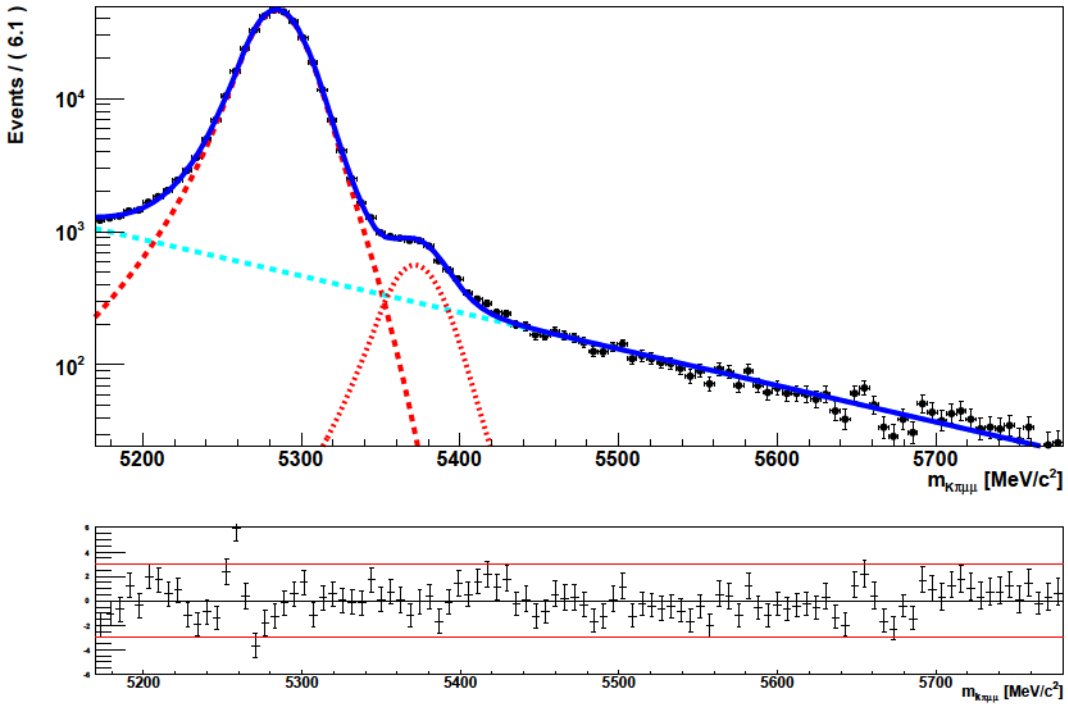
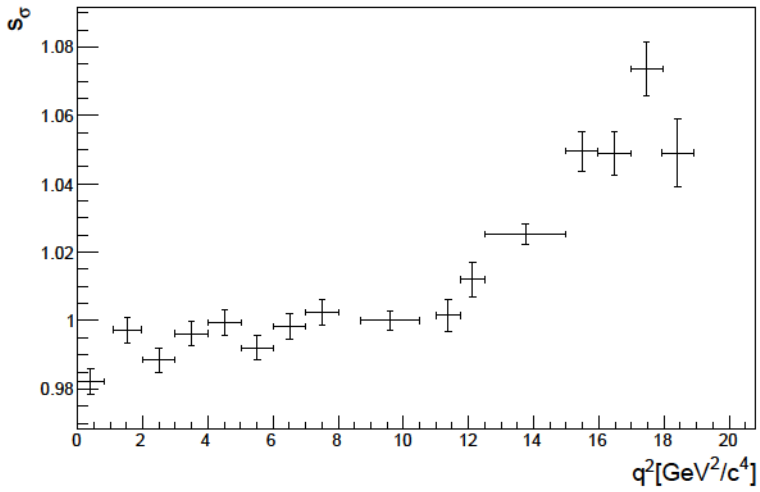


Figure 4.13: Fit to the control channel $B_d \to J/\psi K^{*0}$. The signal component is in red dash-line. the $B_s \to J/\psi \overline{K^{*0}}$ component is in red dotted-line and the combinatorial background in cyan.

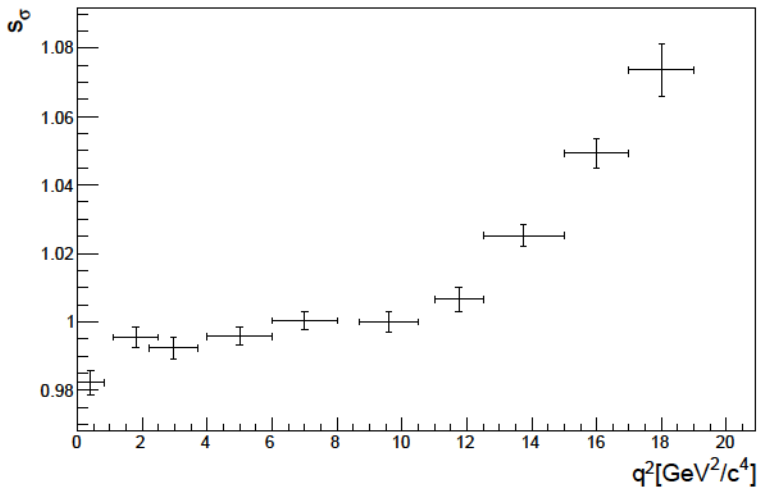
q^2[GeV^2]	Scaling factor
[0.1, 0.98]	0.982
[1.1, 2.0]	0.997
[2.0, 3.0]	0.989
[3.0, 4.0]	0.996
[4.0, 5.0]	0.999
[5.0, 6.0]	0.992
[6.0, 7.0]	0.998
[7.0, 8.0]	1.003
[11.0, 11.75]	1.002
[11.75, 12.5]	1.001
[15.0, 16.0]	1.050
[16.0, 17.0]	1.050
[17.0, 18.0]	1.074
[18.0, 19.0]	1.049

q^2[GeV^2]	Scaling factor
[0.1, 0.98]	0.982
[1.1, 2.5]	0.996
[2.5, 4.0]	0.992
[4.0, 6.0]	0.996
[6.0, 8.0]	1.000
[11.0, 12.5]	1.007
[15.0, 17.0]	1.049
[17.0, 19.0]	1.074

Table 4.5: Scale factor s_σ for the 2 GeV^2 q^2 binning scheme (left) and 1 GeV^2 binning scheme (right).

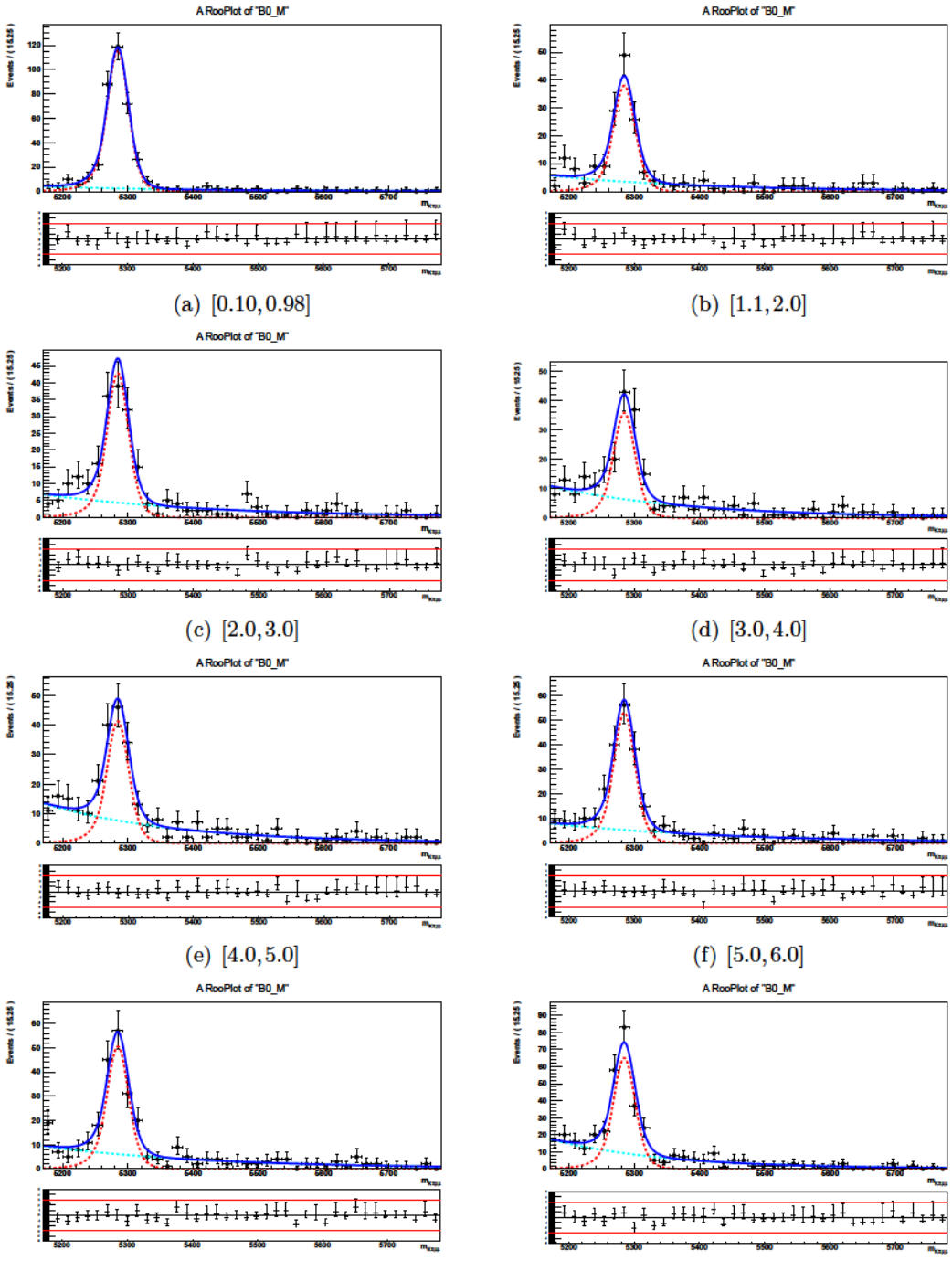


(a) 1 GeV^2 bins



(b) 2 GeV^2 bins

Figure 4.14: Scale factor s_σ in the 1 GeV^2 binning scheme (a) and 2 GeV^2 q^2 binning scheme (b).



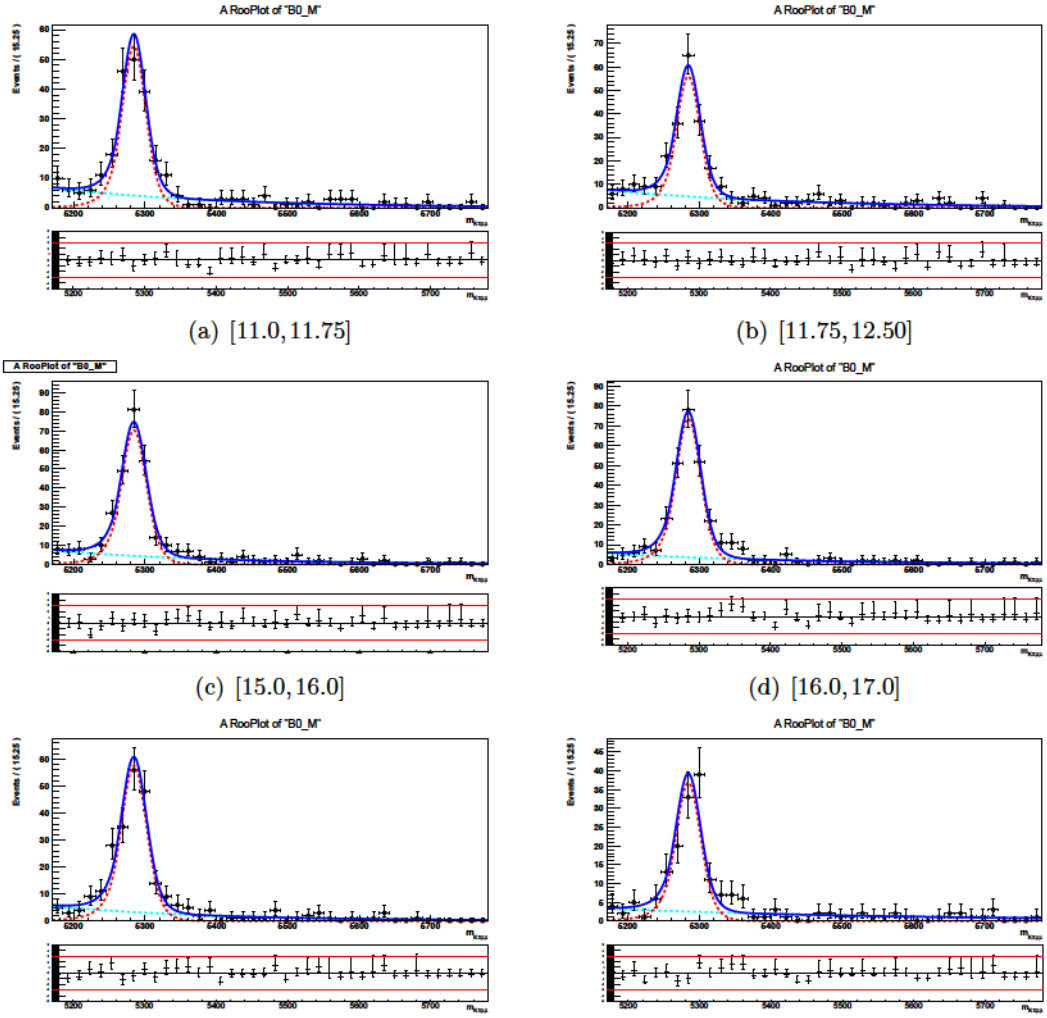
(a) [0.10, 0.98] (b) [1.1, 2.0]

(c) [2.0, 3.0] (d) [3.0, 4.0]

(e) [4.0, 5.0] (f) [5.0, 6.0]

(g) [6.0, 7.0] (h) [7.0, 8.0]

Figure 4.15: The K$\pi\mu\mu$ invariant mass distribution of $B_d \to K^{*0}\mu\mu$ candidates in the first 8 bins in the 1 GeV$^2/c^4$ q^2 binning scheme.

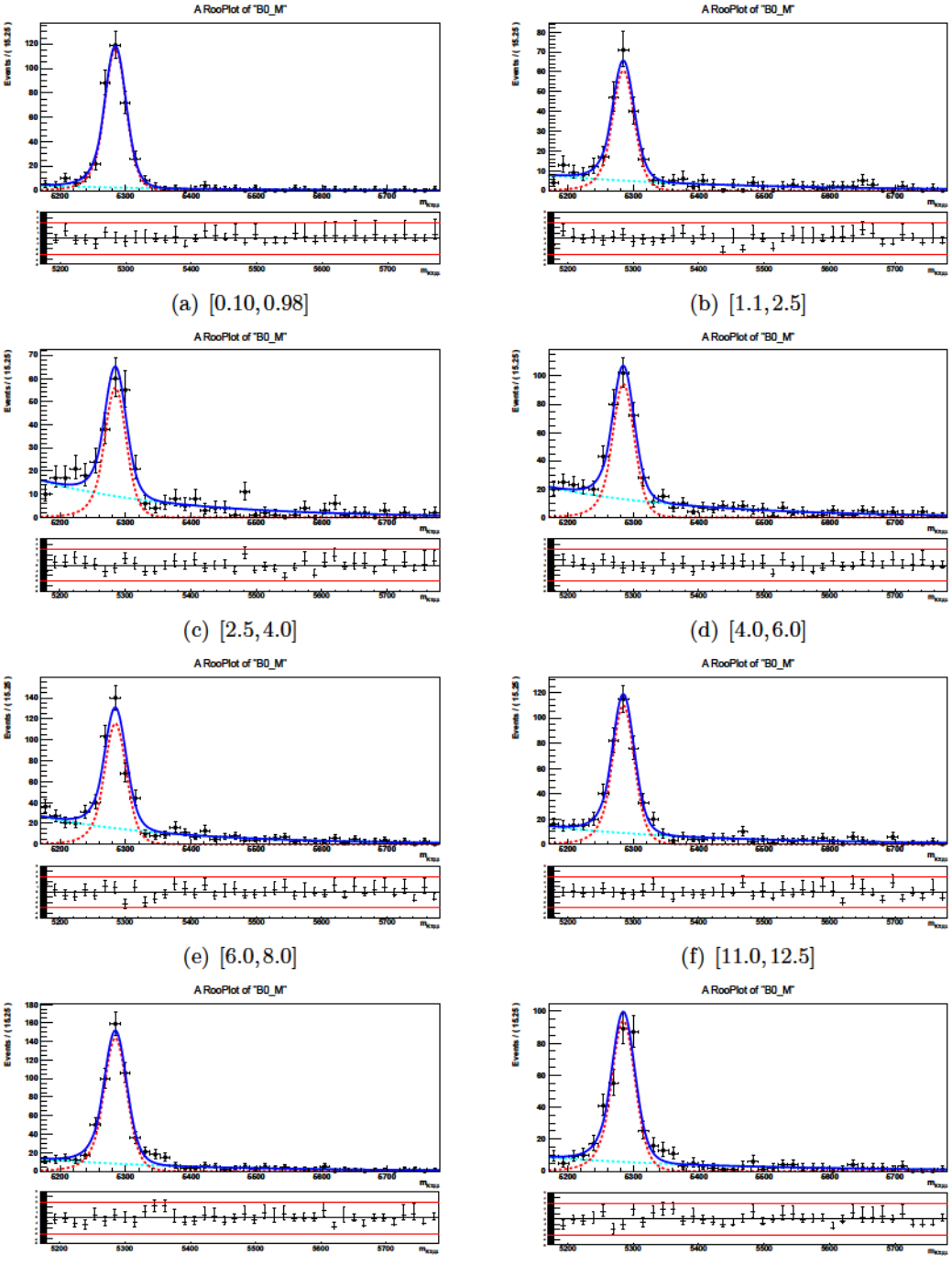


(a) $[11.0, 11.75]$ (b) $[11.75, 12.50]$

(c) $[15.0, 16.0]$ (d) $[16.0, 17.0]$

(e) $[17.0, 18.0]$ (f) $[18.0, 19.0]$

Figure 4.16: The $K^+\pi^-\mu^+\mu^-$ invariant mass distribution of $B^0 \to K^{*0}\mu^+\mu^-$ candidates in the last 6 bins in the $1\,\mathrm{GeV}^2/c^4$ q^2 binning scheme.



(a) $[0.10, 0.98]$ (b) $[1.1, 2.5]$

(c) $[2.5, 4.0]$ (d) $[4.0, 6.0]$

(e) $[6.0, 8.0]$ (f) $[11.0, 12.5]$

(g) $[19.0, 17.0]$ (h) $[17.0, 19.0]$

Figure 4.17: The $K^+\pi^-\mu^+\mu^-$ invariant mass distribution of $B^0 \to K^{*0}\mu^+\mu^-$ candidates in the first 8 bins in the $2\,\mathrm{GeV}^2/c^4$ q^2 binning scheme

.

$q^2[\mathrm{GeV}^2]$	signal yield	background yield
$[0.1, 0.98]$	339.1 ± 19.6	58.9 ± 10.3
$[1.1, 2.0]$	113.0 ± 12.2	82.0 ± 10.9
$[2.0, 3.0]$	126.3 ± 13.3	103.7 ± 12.4
$[3.0, 4.0]$	106.4 ± 12.8	144.6 ± 14.2
$[4.0, 5.0]$	123.1 ± 13.9	169.9 ± 15.5
$[5.0, 6.0]$	156.6 ± 14.6	130.4 ± 13.7
$[6.0, 7.0]$	150.3 ± 14.5	146.7 ± 14.4
$[7.0, 8.0]$	194.2 ± 16.7	197.8 ± 16.8
$[11.0, 11.75]$	162.5 ± 14.7	96.5 ± 12.2
$[11.75, 12.5]$	166.9 ± 15.0	116.1 ± 13.2
$[15.0, 16.0]$	219.5 ± 16.7	102.5 ± 12.8
$[16.0, 17.0]$	229.8 ± 16.9	87.1 ± 12.0
$[17.0, 18.0]$	184.1 ± 15.4	75.9 ± 11.5
$[18.0, 19.0]$	114.9 ± 12.2	69.1 ± 10.2

$q^2[\mathrm{GeV}^2]$	signal yield	background yield
$[0.1, 0.98]$	339.1 ± 19.6	58.9 ± 10.3
$[1.1, 2.5]$	179.7 ± 15.4	124.4 ± 13.5
$[2.5, 4.0]$	165.4 ± 15.9	206.6 ± 17.1
$[4.0, 6.0]$	279.5 ± 20.2	300.4 ± 20.7
$[6.0, 8.0]$	344.3 ± 22.1	344.8 ± 22.1
$[11.0, 12.5]$	329.8 ± 21.0	212.1 ± 18.0
$[15.0, 17.0]$	449.2 ± 23.8	189.8 ± 17.5
$[17.0, 19.0]$	299.9 ± 19.8	144.1 ± 15.3

Table 4.6: Signal and background yields with the two q^2 binning schemes: (top) $1\,\mathrm{GeV}^2/c^4$ and (bottom) $2\,\mathrm{GeV}^2/c^4$.

4.6 Contraining the S-wave contribution using the $m_{K\pi}$ distribution

Including the S-wave contribution results in a reduction of sensitivity to the physically interesting P-wave observables. This is because, according to Eq. 1.65, all P-wave parameters are scaled by the factor $(1-F_S)$ which is not known a priori. Neglecting the S-wave in the fit and correcting the P-wave parameters using F_S from the dedicated S-wave analysis in preparation is problematic since it partially uses the same data distributions ($m_{K\pi\mu\mu}$ and $\cos\theta_K$). A possibility to circumvent these difficulties is to include the $m_{K\pi}$ projection in a simultaneous fit. Since the P-wave is peaking in $m_{K\pi}$ while the S-wave contribution is relatively flat, this gives an additional constraint on F_S and therefore also allows a better determination of the P-wave observables. Ref. [8] gives details on the dependence of the decay amplitudes on $m_{K\pi}$. To parameterise the $m_{K\pi}$ dependence of the P-wave a Breit-Wigner distribution is used

$$\begin{aligned}\mathcal{A}_P(m_{K\pi}) = \sqrt{pq} \times B'_{L_B}(p,p_0,d)\left(\frac{p}{m_B}\right)^{L_B} \times B'_{L_{K^*}}(q,q_0,d)\left(\frac{q}{m_{K\pi}}\right)^{L_{K^*}} \\ \times \frac{1}{m^2_{K^*} - m^2_{K\pi} - i m_{K^*}\Gamma(m_{K\pi})},\end{aligned} \tag{4.17}$$

where p (q) denotes the K^{*0} (K^+) momentum in the B^0 (K^{*0}) rest frame, p_0 (q_0) is the corresponding quantity at the resonance peak. L_B (L_{K^*}) is the orbital angular momentmum and B'_{L_B} ($B'_{L_{K^*}}$) the Blatt-Weisskopf function given in Ref. [9]. For the S-wave component the LASS parameterisation [10] is used

$$\begin{aligned}\mathcal{A}_S(m_{K\pi}) = \sqrt{pq} \times B'_{L_B}(p,p_0,d)\left(\frac{p}{m_B}\right)^{L_B} \times B'_{L_{K_0^*}}(q,q_0,d)\left(\frac{q}{m_{K\pi}}\right)^{L_{K_0^*}} \\ \times \left(\frac{1}{\cot\delta_B - i} + e^{2i\delta_B}\frac{1}{\cot\delta_R - i}\right),\end{aligned} \tag{4.18}$$

where $\cot\delta_B = \frac{1}{aq} + \frac{1}{2}rq$ and $\cot\delta_R = (m^2_{K_0^*} - m^2_{K\pi})/(m_{K_0^*}\Gamma_0(m_{K\pi}))$. Accounting for the $m_{K\pi}$ dependence, Eq. 1.65, integrated over the three decay angles $\cos\theta_l$, $\cos\theta_K$ and ϕ, becomes

$$\begin{aligned}\left.\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}m_{K\pi}}\right|_{\mathrm{S+P}} &= (1-F_S)\sum_{i=1}^{9}\tfrac{9}{32\pi}\xi_i S_i^{(s,c)}|\mathcal{A}'_P(m_{K\pi})|^2 \\ &+ \tfrac{3}{16\pi}\big[F_S\xi_{F_S}|\mathcal{A}'_S(m_{K\pi})|^2 \\ &\quad + (S_{S1}\xi_{S1} + S_{S2}\xi_{S2} + S_{S3}\xi_{S3})\,\Re\left(\mathcal{A}'_S(m_{K\pi})\mathcal{A}'^*_P(m_{K\pi})\right) \\ &\quad + (S_{S4}\xi_{S4} + S_{S5}\xi_{S5})\,\Im\left(\mathcal{A}'_S(m_{K\pi})\mathcal{A}'^*_P(m_{K\pi})\right)\big]\end{aligned} \tag{4.19}$$

where $\xi_{(S)i}$ denote the angular integrals $\xi_{(S)i} = \int \epsilon(\cos\theta_l,\cos\theta_K,\phi) f_{(S)i}(\cos\theta_l,\cos\theta_K,\phi)\mathrm{d}\vec{\Omega}$ and the amplitudes are appropriately normalised according to

$$\mathcal{A}'_P(m_{K\pi}) = \frac{\mathcal{A}_P(m_{K\pi})}{\sqrt{\int_{795.9\,\mathrm{MeV}/c^2}^{995.9\,\mathrm{MeV}/c^2}|\mathcal{A}_P(m_{K\pi})|^2\mathrm{d}m_{K\pi}}},$$

$$\mathcal{A}'_S(m_{K\pi}) = \frac{\mathcal{A}_S(m_{K\pi})}{\sqrt{\int_{795.9\,\mathrm{MeV}/c^2}^{995.9\,\mathrm{MeV}/c^2}|\mathcal{A}_S(m_{K\pi})|^2\mathrm{d}m_{K\pi}}}.$$

In the case of flat acceptance the integrated terms $\xi_{S1\ldots5}$ evaluate to $\xi_{S1\ldots5} = 0$ such that the interference terms drop out. For the nominal acceptance these terms are of the order of a few percent and are included for completeness.

The simultaneous fit of the angles and the $m_{K\pi}$ projection is tested using the control decay $B^0 \to J/\psi K^{*0}$. Table 4.7 gives the results of a fit of the full $3\,\text{fb}^{-1}$ data sample in the $m_{K\pi}$ mass region $\pm 100\,\text{MeV}$ around the K^{*0} mass. Fig. 4.18 shows the corresponding projections on the decay angles, $m_{K\pi\mu\mu}$ and $m_{K\pi}$.

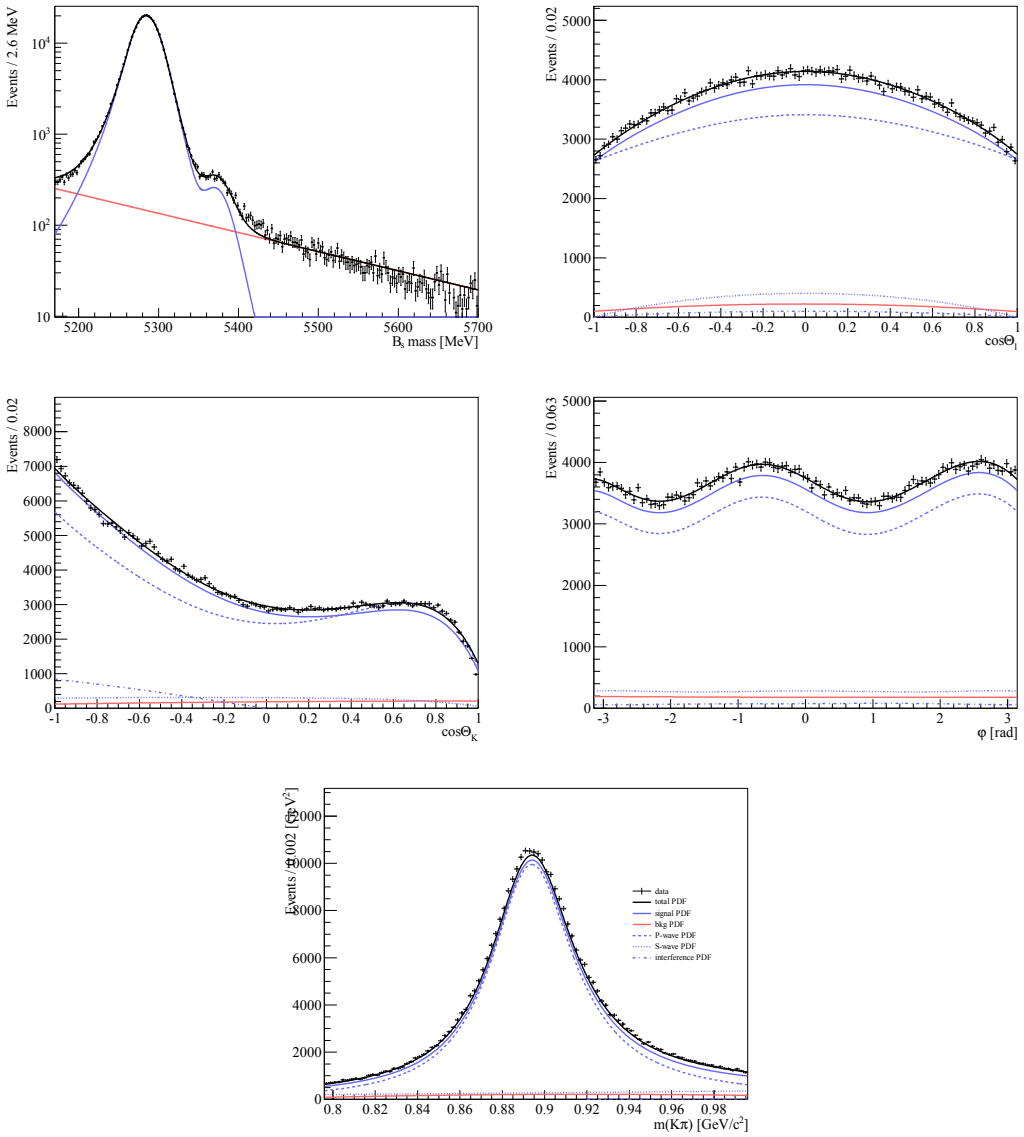


Figure 4.18: Angular, $m_{K\pi\mu\mu}$ and $m_{K\pi}$ projections after the fit of the full $B^0 \to J/\psi K^{*0}$ data sample. The fit is performed as described in Sec. 4.6, simultaneously in the decay angles and $m_{K\pi\mu\mu}$, and $m_{K\pi}$, in the $m_{K\pi}$ mass range $[795.9, 995.9]\,\text{MeV}$.

Table 4.7: Result of the simultaneous fit of the decay angles, $m_{K\pi\mu\mu}$ and $m_{K\pi}$ for the full $B^0 \to J/\psi\, K^{*0}$ data sample.

parameter	value
S_{1s}	0.3309 ± 0.0010
S_3	0.0015 ± 0.0018
S_4	-0.2757 ± 0.0020
S_5	-0.0023 ± 0.0019
S_{6s}	0.0017 ± 0.0016
S_7	0.0008 ± 0.0020
S_8	-0.0503 ± 0.0020
S_9	-0.0866 ± 0.0019
F_S	0.0793 ± 0.0028
S_{S1}	-0.2286 ± 0.0034
S_{S2}	0.0009 ± 0.0022
S_{S3}	0.0025 ± 0.0021
S_{S4}	0.0013 ± 0.0021
S_{S5}	-0.0653 ± 0.0023

REFERENCES

[1] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay* $B^0 \to K^{*0}\mu^+\mu^-$, JHEP **08** (2013) 131, arXiv:1304.6325.

[2] LHCb collaboration, R. Aaij *et al.*, *Measurement of form-factor-independent observables in the decay* $B^0 \to K^{*0}\mu^+\mu^-$, Phys. Rev. Lett. **111** (2013) 191801, arXiv:1308.1707.

[3] BABAR Collaboration, B. Aubert *et al.*, *Measurements of branching fractions, rate asymmetries, and angular distributions in the rare decays* $B \to K\ell^+\ell^-$ *and* $B \to K^*\ell^+\ell^-$, Phys. Rev. D **73** (2006) 092001.

[4] Belle Collaboration, J.-T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for* $B \to K^{(*)}l^+l^-$,Phys. Rev. Lett. **103** (2009) 171801.

[5] ATLAS collaboration, G. A. et al, *Measurement of the inclusive isolated prompt photon cross section in pp collisions at* $\sqrt{7}$ *TeV with the ATLAS detector*, ArXiv Physics e-prints (2010) arXiv:1012.4389.

[6] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, ArXiv Physics e-prints (2003) arXiv:physics/0306116.

[7] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*,.

[8] D. Becirevic and A. Tayduganov, *Impact of* $B \to K_0^*\ell^+\ell^-$ *on the New Physics search in* $B \to K^*\ell^+\ell^-$ *decay*, Nucl. Phys. **B868** (2013) 368, arXiv:1207.4004.

[9] J. Bressieux, G. Cowan, M. Kreps, O. Schneider, T. Skwarnicki, Q. Wenbin, *Evidence for the resonant character of the* $Z(4430)^- \to \psi(2S)\pi^-$ *mass peak observed in* $B^0 \to \psi(2S)K^+\pi^-$ *decays, and determination of the* $Z(4430)^-$ *spin-parity* , LHCb-ANA-2013-053, 2013.

[10] D. Aston *et al.*, *A Study of K- pi+ Scattering in the Reaction K- p —> K- pi+ n at 11-GeV/c*, Nucl. Phys. **B296** (1988) 493.

Chapter 5

Results and systematic uncertainties

Sommaire

In the first part of this chapter a brief description of the different sources of systematics uncertainties affecting the measurements of the angular observables is provided (Section 5.1). Some of these uncertainties are related to the measurement of the analysis acceptance, as described in Section 5.1.1. Uncertainties associate to the modeling of the backgrounds are reported in Section 5.1.2 and mainly concerns the peaking backgrounds estimation and the background angular distributions. Effects related to the signal invariant mass modeling (Section 5.1.3) and invariant mass of the $K\pi$ system (Section 5.1.4) are also taken into account as well as possible

asymmetries in the production of B mesons and in the detection (Sections 5.1.5 and 5.1.6 respectively). A summary of the different systematic uncertainties for each set of observables is provided in Section 5.1.7.

In the second part, the results of the measurement of the angular observables, determined using the angular maximum likelihood fit described in Section. 4.1 on the 3 fb^{-1} of data collected during the first run of LHC are are given in the Section 5.2, together with a short overview of possible theoretical interpretations (Section 5.3) and of future experimental tests (Section 5.4) that could help in understanding these results.

5.1 Systematics

Systematics uncertainties are computed as the variation between the results of fits on high statistics pseudo-experiments using once the nominal and once the systematically varied models. Each toy uses a larger number of events compared with the data in order to eliminate any statistical uncertainty in the systematic error evaluation.

5.1.1 Systematic errors concerning the acceptance

5.1.1.1 Statistical uncertainty of the four-dimensional acceptance

The four-dimensional acceptance described in Section 4.4 relies on the determination of the Legendre coefficients from a sample of simulated $B^0 \to K^{*0}\mu^+\mu^-$ phase-space events. To determine the effect of the limited size of the simulated sample, the covariance matrix of the coefficients is determined alongside their numerical values. High statistics toy studies are then performed, where the events are generated with an acceptance that is varied according to the (inverse) covariance matrix. These simulated events are then fit using both the varied and the nominal acceptance. Figure 5.1 gives the distributions for 500 toy experiments for the q^2 bin in the range $0.1 < q^2 < 1.0\,\mathrm{GeV}^2/c^4$. The observed deviations of the parameters are fit using Gaussian functions. The distributions are centered around 0, their widths are used as systematic uncertainties due to the statistical uncertainty of the acceptance. The Appendix D.1, Table D.1 gives the systematic uncertainties for all q^2 bins. They are negligible compared to the expected statistical uncertainties.

5.1.1.2 Differences between data and simulation

The determination of the acceptance relies on an accurate simulation of the signal decay $B^0 \to K^{*0}\mu^+\mu^-$. The control decay $B^0 \to J/\psi K^{*0}$ is used to cross-check if the distributions in data are reproduced properly. Ref. [1] describes the procedure used to correct the unsatisfactory simulation of the transverse momentum of the signal B^0, as well as the B^0 vertex χ^2 and the track multiplicity in the event. The effect of these corrections on the acceptance is evaluated by redetermining the acceptance correction without the reweighting. The results are given in Tables D.2, D.3 and D.4. All deviations seen are negligible.

In addition, there are small differences for the kinematic variables of the B^0 daughter particles. Figure 5.2 shows the (transverse) momentum for the signal kaon and pion for both truth-matched simulated events as well as $B^0 \to J/\psi K^{*0}$ events from data. The distributions for data are extracted using the *sWeighting* technique. To minimize the influence of pollution from an S-wave component which is not simulated in data, the window for invariant mass of the $K^+\pi^-$ system is reduced from the nominal $\pm 100\,\mathrm{MeV}/c^2$ to $\pm 20\,\mathrm{MeV}/c^2$. From the two-dimensional distributions of K^+ and π^- in data and simulation a correction factor is determined. This

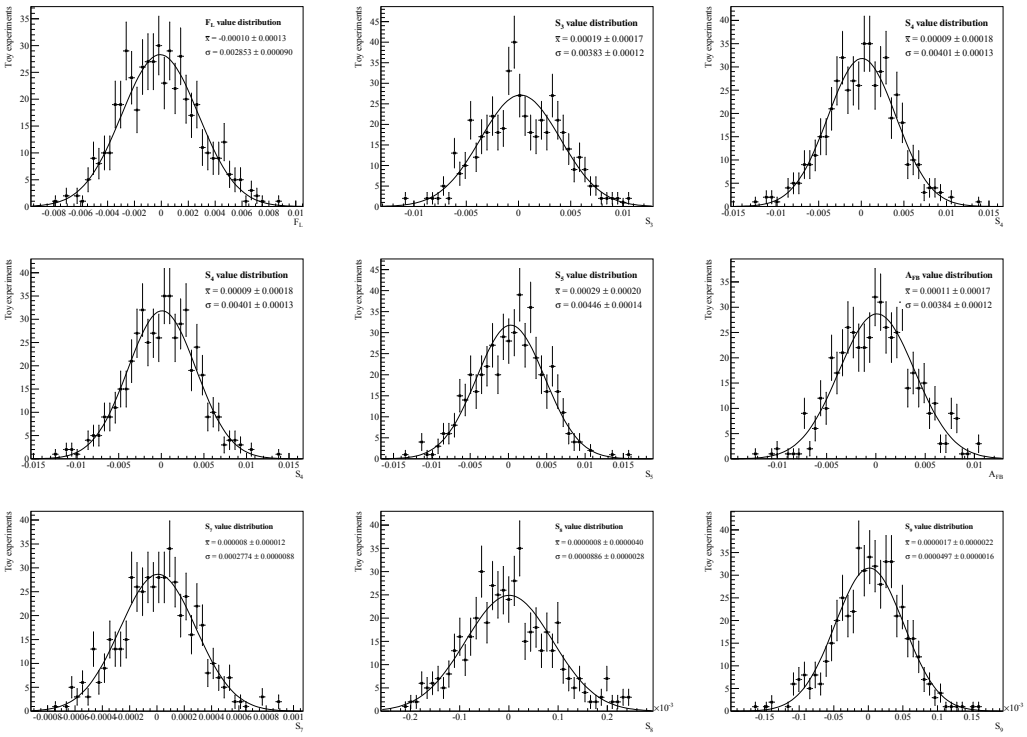


Figure 5.1: Distributions of the deviations of the observables from toy experiments for the first q^2 bin in the range $0.1 < q^2 < 1.0\,\mathrm{GeV}^2/c^4$. Events are generated with an acceptance varied according to its statistical uncertainty and then fit back using the nominal acceptance.

correction factor, depending on the particles momentum and transverse momentum is shown in Figure 5.3. The systematic uncertainty from the modeling of the signal decay is then evaluated using toy studies where the acceptance is redetermined using the reweightings. Tables D.5 and D.6 give the resulting systematic uncertainties for the reweighting of both kaon and pion. While the reweighting of the kaon has a negligible effect, there is an, albeit small, systematic uncertainty for the reweighting of the pion.

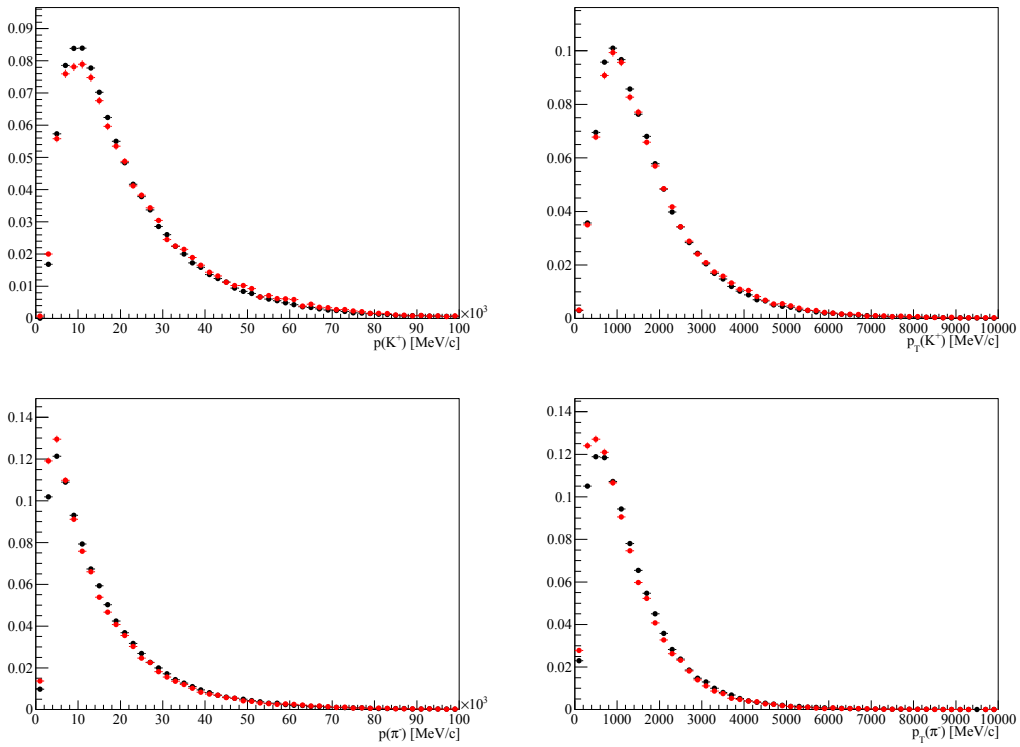


Figure 5.2: Distribution of p and p_T for (up) K^+ and (down) π^- for *sWeighted* $B^0 \to J/\psi\, K^{*0}$ candidates from data (black) and truthmatched $B^0 \to J/\psi\, K^{*0}$ decays from simulation (red).

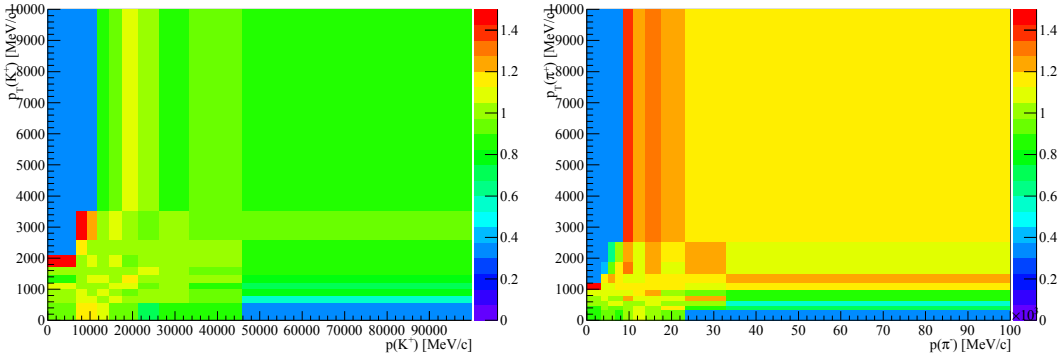


Figure 5.3: Correction factor for simulated events depending on p and p_T for (left) K^+ and (right) π^-.

5.1.1.3 Impact of the fixed q^2 value per bin in the four-dimensional acceptance

Since a weighted fit of the data is less stable in narrow bins[1], given their lower number of events, the efficiency has been included in the PDF in the fit. The PDF used to fit the observables is in itself not q^2 dependent, therefore the four-dimensional efficiency is evaluated for a fixed q^2. The nominal setting is to use the mean of the q^2 in the bin to determine the efficiency used for the specific bin.

To evaluate a possible systematic bias from this choice, toy experiments are used to determine the deviation in the observables when fixing q^2 at the mean value in the bin, q^2_{mean}, respect to the value at the edge of the bin. More explicitly, the toys use the acceptance at $q^2_{\text{mean}} \pm \frac{3}{4}\Delta q^2$, where the q^2 bin is given by $[q^2_{\text{mean}} - \Delta q^2, q^2_{\text{mean}} + \Delta q^2]$. The largest deviation is taken as systematic uncertainty and given in Table. D.7 for all observables and q^2 bins. While this approach likely overestimates the systematic effect, the resulting systematic uncertainties are small compared to the expected statistical errors.

5.1.1.4 Higher order acceptance model

There is some choice in the maximum order of the Legendre polynomials used to model the four-dimensional acceptance. While higher orders generally will be able to describe details in the acceptance better, more coefficients also will lead to higher computational requirements. Therefore, the lowest order polynomials that describe the acceptance sufficiently well has been chosen. A higher order parametrisation is used to determine a systematic uncertainty on this choice.

The nominal choice is to include Legendre polynomials of order four and lower for $\cos\theta_l$, order five for $\cos\theta_k$ and q^2, and order six for the angle ϕ as described in Section 4.4. In addition, the acceptance is assumed to be even in ϕ, resulting in only non-zero coefficients of even order for these polynomials. The projections of the four-dimensional acceptance determined with these settings provide a very good description of the angles $\cos\theta_l$ and ϕ (See Figure 4.7). Small deviations are observed for low q^2 and large $\cos\theta_K$. To estimate the effect of these imperfections on the angular observables, we determine an acceptance including higher order polynomials for the description of $\cos\theta_k$ and q^2, choosing a maximal order of seven for both. Table D.8 gives the result of the angular fit of the control decay $B^0 \to J/\psi\, K^{*0}$ using this higher order acceptance and for comparison the nominal result. No deviation of a significant size for the angular analysis of the signal decay is seen. This gives again confidence in the choice of the acceptance description.

To determine the systematic uncertainties properly, high statistics toys are performed, where events are generated using the higher order acceptance model and fit with the nominal one. The resulting deviations are given in Table D.9, and they are negligible for all bins.

5.1.2 Background related systematics

5.1.2.1 Systematic error on the peaking backgrounds

Several peaking backgrounds are able to mimic the signal decay. An overview is given in Table 5.1 taken from [1], where the peaking background processes are discussed in more detail. To determine the effect of neglecting the peaking backgrounds in the angular analysis, high statistics toy studies are performed. In addition to the signal and combinatorial background component, peaking background events are added according to their expected fraction. The

[1] The unfolding can be used for the larger q^2 bins $1 < q^2 < 6\,\text{GeV}^2/c^4$ and $15 < q^2 < 19\,\text{GeV}^2/c^4$.

deviation of the fitted angular observables from their nominal values when neglecting these peaking background events in the fit are then taken as systematic uncertainties.

	after preselection, before vetoes		after vetoes and selection	
Channel	Estimated events	% signal	Estimated events	% signal
$\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-$	$(1.0 \pm 0.5) \times 10^3$	19 ± 8	51 ± 25	1.0 ± 0.4
$\Lambda_b^0 \to pK^-\mu^+\mu^-$	$(1.0 \pm 0.5) \times 10^2$	1.9 ± 0.8	5.7 ± 2.8	0.11 ± 0.05
$B^+ \to K^+\mu^+\mu^-$	28 ± 7	0.55 ± 0.06	1.6 ± 0.5	0.031 ± 0.006
$B_s^0 \to \phi\mu^+\mu^-$	$(3.2 \pm 1.3) \times 10^2$	6.2 ± 2.1	17 ± 7	0.33 ± 0.12
signal swaps	$(3.6 \pm 0.9) \times 10^2$	6.9 ± 0.6	33 ± 9	0.64 ± 0.06
$B^0 \to J/\psi K^{*0}$ swaps	$(1.3 \pm 0.4) \times 10^2$	2.6 ± 0.4	2.7 ± 2.8	0.05 ± 0.05
$B^0 \to J/\psi K^{*0}$	70 ± 22	1.35 ± 0.28	59 ± 19	1.14 ± 0.26
$B^+ \to K^{*+}\mu^+\mu^-$	0	0	0	0

Table 5.1: Estimated yields, and percentage relative to estimated signal yield, of peaking background events before and after the vetoes. The dominant uncertainty contributing to these numbers is in $\sigma_{b\bar{b}}$ and the estimate of $\mathcal{B}(\Lambda_b^0 \to \Lambda^*(1520)^0\mu^+\mu^-)$.

The distributions of the peaking background events in reconstructed B^0 mass, decay angles and q^2 are taken from data. To select these peaking background events, specific selections are applied. The explicit vetoes against the peaking backgrounds are removed and the criteria listed in Table 5.2 are applied instead. Since the nominal BDT used to suppress combinatorial background includes particle identification criteria, the nominal BDT cut is removed. Instead, a new BDT, trained without particle identification criteria, is applied to remove combinatorial backgrounds.

mode	selection criteria
$\Lambda_b^0 \to pK\mu^+\mu^-$	$\texttt{ProbNNk}(K) > 0.3$
	$\texttt{ProbNNp}(\pi) > 0.3$
	$\lvert m_{pK\mu\mu} - m_{\Lambda_b^0}\rvert < 50\,\text{MeV}/c^2$
$B_s^0 \to \phi\mu^+\mu^-$	$\texttt{ProbNNk}(K,\pi) > 0.3$
	$\lvert m_{KK} - m_{\phi}\rvert < 20\,\text{MeV}/c^2$
	$\lvert m_{KK\mu\mu} - m_{B_s^0}\rvert < 50\,\text{MeV}/c^2$
$B^0 \to \pi^+\pi^-\mu^+\mu^-$	$\texttt{ProbNNpi}(K,\pi) > 0.3$
	$\texttt{ProbNNk}(K,\pi) < 0.1$
	$\lvert m_{\pi\pi\mu\mu} - m_{B^0}\rvert < 50\,\text{MeV}/c^2$

Table 5.2: Particle identification criteria and mass ranges to explicitly select specific peaking backgrounds.

The selected peaking backgrounds $\Lambda_b^0 \to pK\mu^+\mu^-$, $B_s^0 \to \phi\mu^+\mu^-$ and $B^0 \to \pi^+\pi^-\mu^+\mu^-$, as well as $K\pi$ swapped $B^0 \to K^{*0}\mu^+\mu^-$ events, are given in Figure D.1. Here, the standard charmonium vetoes are applied. In addition, Figure D.2 gives the corresponding high statistics charmonium modes $\Lambda_b^0 \to J/\psi pK$, $B_s^0 \to J/\psi\phi$, $B^0 \to J/\psi\pi^+\pi^-$ and $B^0 \to J/\psi K^{*0}$ swaps, where q^2 is in the range $[8,11]\,\text{GeV}^2/c^4$. The selected peaking background yields are 109 ($\Lambda_b^0 \to pK\mu^+\mu^-$), 156 ($B_s^0 \to \phi\mu^+\mu^-$) and 92 ($B^0 \to \pi^+\pi^-\mu^+\mu^-$) events. As expected, the charmonium decays have much larger yields with 9,000 ($\Lambda_b^0 \to J/\psi pK$), 24,100 ($B_s^0 \to J/\psi\phi$) and 9,000 ($B^0 \to J/\psi\pi^+\pi^-$) events.

Two different methods are employed to determine the angular distributions of the peaking background events. The first is to simply sample events randomly from the high statistics $b \to J/\psi X$ decays, using the q^2 distribution of the corresponding rare modes to determine the fraction of background events expected in the different q^2 bins. The second approach is to use a kernel density method to describe the distributions, using only the low statistics $b \to X\mu^+\mu^-$ modes. This method uses Gaussian kernels according to

$$\begin{aligned}\mathcal{P}_{\text{peaking}}(m_{K\pi\mu\mu}, \cos\theta_l, \cos\theta_k, \phi) = \frac{1}{N} \sum_{\text{events}}^{N} \sum_{i=1} & \frac{1}{\sqrt{2\pi}^4 \sigma(\cos\theta_l)\sigma(\cos\theta_k)\sigma(\phi)\sigma(m_{K\pi\mu\mu})} \\ & \times \exp\left[-\frac{(m_{K\pi\mu\mu} - m_{K\pi\mu\mu,i})^2}{2\sigma^2(m_{K\pi\mu\mu})} - \frac{(\cos\theta_l - \cos\theta_{l,i})^2}{2\sigma^2(\cos\theta_l)} \right] \\ & \times \exp\left[-\frac{(\cos\theta_k - \cos\theta_{K,i})^2}{2\sigma^2(\cos\theta_k)} - \frac{(\phi - \phi_i)^2}{2\sigma^2(\phi)} \right], \end{aligned} \tag{5.1}$$

where $\sigma(\cos\theta_l) = \sigma(\cos\theta_k) = 0.2$, $\sigma(\phi) = \pi/5\,\text{rad}$ and $\sigma(m_{K\pi\mu\mu}) = 10.6\,\text{MeV}/c^2$. Events near the borders of the $\cos\theta_l$ and $\cos\theta_k$ distributions are handled by folding back the PDF. Figure D.3 shows the angular distributions for $b \to X\mu^+\mu^-$ decays in black. Overlayed are the angular distributions of the $b \to J/\psi X$ decays in blue and the results from the kernel method in red. The results from the kernel method follow the data smoothly, the angular distributions from the charmonium modes seem to be statistically compatible with the rare decays. The most interesting feature is certainly the $\cos\theta_k$ dependence of the $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to J/\psi\,\phi$ decays which strongly peak towards $\cos\theta_k = -1$. This is due to the mass of the ϕ resonance being just above the K^+K^- threshold.

In addition to the peaking backgrounds from misidentified $\Lambda_b^0 \to pK\mu^+\mu^-$, $B_s^0 \to K^+K^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-\mu^+\mu^-$ decays, there are 2% of $B^0 \to K^+\pi^-_{\text{rnd}}\mu^+\mu^-$ decays originating from $B^0 \to K^{*0}\mu^+\mu^-$ decays, where the π^- was replaced by a random pion in the event. The distributions of this background source are modelled using $B^0 \to K^{*+}(\to K^+\pi^0)\mu^+\mu^-$ decays.

The resulting deviations from high statistics toys containing the appropriate fraction of $\Lambda_b^0 \to pK\mu^+\mu^-$, $B_s^0 \to K^+K^-\mu^+\mu^-$ and $B^0 \to K^+\pi^-\mu^+\mu^-$, as well as $B^0 \to K^+\pi^-_{\text{rnd}}\mu^+\mu^-$ peaking background events sampled from the charmonium decays are given in the Table D.10. The corresponding results from the kernel method are given in Table D.11. The peaking background systematics uncertainty is defined as the biggest uncertainty between the kernel density method and the high statistics charmonium mode for each observables and each q^2 bins.

5.1.2.2 Angular background modeling

The nominal background parametrisation uses Chebyshev polynomials of second order and lower to describe the decay angles. To estimate the systematic effect of this choice of angular background model, the high mass sideband ($m_{K\pi\mu\mu} \in [5355, 5700]\,\mathrm{MeV}/c^2$) is fit with Chebyshev polynomials of forth order and lower instead. To have enough combinatorial background events to fit the Chebyshev coefficients, the BDT requirement is removed for the fit. Figure D.4 shows the fitted angular distributions in bins of q^2. To determine the systematic effect of only fitting the background with polynomials of order two and lower, high statistics toy MC are used. The toy MC is generated using the forth order angular background description using the nominal signal fraction. Then the toys are fitted once with the fourth order and once with the second order angular background description. The observed difference is taken as systematic uncertainty and is given in Table D.12. The systematic effect is negligible.

5.1.3 Signal mass modeling

To determine the systematic error related to the choice of signal mass model, described in Section 4.5, a double Gaussian is used as alternative model. The parameters of the double Gaussian are determined from a fit to $B^0 \to J/\psi K^{*0}$ events. A high statistics toy MC is then generated using the double Gaussian mass model and fitted twice, once using the double Gaussian and once using the nominal Crystal Ball parametrisation. The observed difference is given in Table D.13 and used as systematic uncertainty.

5.1.4 $m_{K\pi}$ related systematic uncertainties

The nominal fit uses the $m_{K\pi}$ distribution to constrain F_S as described in Section 4.6. Three possible sources of systematic uncertainties connected to the $m_{K\pi}$ distribution are studied: the parametrisation of the S-wave component, the parametrisation of the background in $m_{K\pi}$, and the effect of an $m_{K\pi}$ dependent efficiency which is neglected in the nominal fit.

To evaluate the systematic uncertainty due to using the LASS shape as nominal model for the S-wave contribution, the effect of using the ISOBAR model instead is evaluated. The ISOBAR model consists of the sum of two amplitudes modelling the f_{800} and the $K_0^{*0}(1430)$ mesons,

$$\mathcal{A}_{\mathrm{ISOBAR}}(m_{K\pi}) = |r_{f_{800}}| e^{i \arg \delta_{f_{800}}} \mathcal{A}_{f_{800}}(m_{K\pi}) + (1 - |r_{f_{800}}|) \mathcal{A}_{K_0^{*0}(1430)}(m_{K\pi}),$$

where $\mathcal{A}_{f_{800}}(m_{K\pi})$ and $\mathcal{A}_{K_0^{*0}(1430)}(m_{K\pi})$ are Breit-Wigner amplitudes as in Eq. 4.17. The masses and widths of the resonances are set to $m(f_{800}) = 682\,\mathrm{MeV}/c^2$ and $\Gamma(f_{800}) = 547\,\mathrm{MeV}/c^2$ for the f_{800} contribution and $m(K_0^{*0}(1430)) = 1.425\,\mathrm{GeV}/c^2$ and $\Gamma(K_0^{*0}(1430)) = 0.270\,\mathrm{GeV}/c^2$ for the $K_0^{*0}(1430)$ [2]. The parameters $|r_{800}|$ and δ_{800} are determined from a fit to the $m_{K\pi\mu\mu}$ and $m_{K\pi}$ distributions of $B^0 \to J/\psi K^{*0}$ decays. An high statistics toy Monte Carlo is generated using the ISOBAR model and fit twice, once using the ISOBAR, and once the nominal LASS model. The observed deviations for the angular observables are used as systematic uncertainties and given in Tab. D.14.

To evaluate the systematic uncertainty due to the background parametrisation of $m_{K\pi}$ a first order Chebyshev polynomial is compared to a fourth order parametrisation. The fourth order coefficients are determined from $B^0 \to J/\psi K^{*0}$. High statistics toy Monte Carlo is generated using the fourth order parametrisation and fit using both the fourth order and the nominal first order parametrisation. The observed deviations for the angular observables are used as systematic uncertainties and given in Table. D.15.

For the nominal fit the efficiency over the $m_{K\pi}$ range of the angular analysis, [795.9, 995.9] MeV/c^2, is assumed to be flat. The systematic effect of this assumption is quantified using high statistics toy Monte Carlo, including an additional $m_{K\pi}$ dependent efficiency. This efficiency is parametrised using a linearly rising or falling function with a variation of $\pm 5\%$ at the borders of the $m_{K\pi}$ mass range. The additional efficiency is applied on top of the usual four-dimensional efficiency described in Section 4.4. The angular observables are then determined using the nominal fit, and the largest deviations from the generated values are taken as systematic uncertainties and given in Table. D.16.

5.1.5 Production asymmetry

The production of B^0 and $\overline{B}^0$ mesons is known to be asymmetric at the LHC, due to the non-charge symmetric initial state in pp collision. The production asymmetry $\mathcal{A}_{\text{prod}}$, defined as

$$\mathcal{A}_{\text{prod}} = \frac{N(\overline{B}^0) - N(B^0)}{N(\overline{B}^0) - N(B^0)},$$

is measured to be $(-0.35 \pm 0.76 \pm 0.28)\%$ [3]. This affects both the measured CP asymmetries and the CP-averaged observables, according to

$$A_i^{\text{meas}} = A_i - S_i(\kappa \mathcal{A}_{\text{prod}}),$$
$$S_i^{\text{meas}} = S_i - A_i(\kappa \mathcal{A}_{\text{prod}}),$$

where κ is a dilution factor due to $B^0 - \overline{B}^0$ mixing. For B^0 mixing, κ is calculated via

$$\kappa = \frac{\int_0^\infty \epsilon(t) e^{-\Gamma t} \cos(\Delta m_d t) \mathrm{d}t}{\int_0^\infty \epsilon(t) e^{-\Gamma t} \mathrm{d}t},$$

with $\Gamma = 1/\tau_d = 1/1.519\,\text{ps}^{-1}$ and the mixing frequency $\Delta m_d = 0.510\,\text{ps}^{-1}$ [2]. The decay time dependent efficiency $\epsilon(t)$ is given in Figure. 5.4. The calculation results in a factor of $\kappa = 35.2\%$.

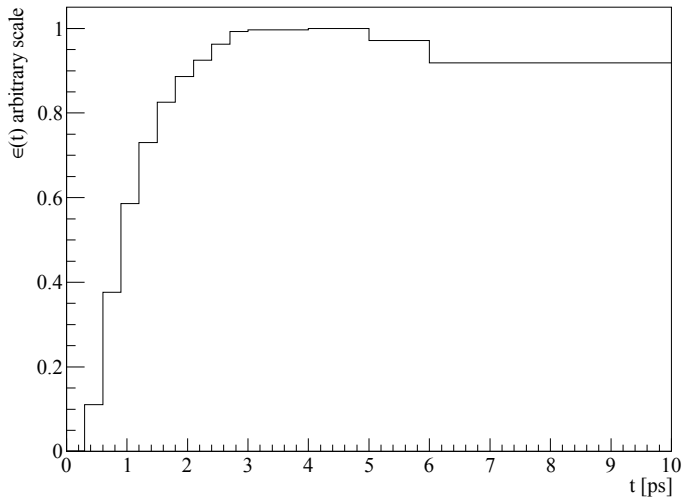


Figure 5.4: The decay time dependent selection efficiency $\epsilon(t)$.

5.1.6 Detection asymmetry

Similarly to the production asymmetry, the measurement can also be affected by the $K^+\pi^-$ detection asymmetry

$$\mathcal{A}_{\text{det}} = \frac{\epsilon(K^+\pi^-) - \epsilon(K^-\pi^+)}{\epsilon(K^+\pi^-) + \epsilon(K^-\pi^+)},$$

according to

$$A_i^{\text{meas}} = A_i - S_i\mathcal{A}_{\text{det}},$$
$$S_i^{\text{meas}} = S_i - A_i\mathcal{A}_{\text{det}}.$$

The $K^+\pi^-$ detection asymmetry is known to be driven by the kaon detection asymmetry, which is momentum dependent. The kaon detection asymmetry, in bins of kaon momentum, is given in Table. 5.3, which is taken from Ref. [4] and was used in Ref. [5]. Since the momentum spectra for the hadrons depend on q^2, the detection asymmetry is determined for all q^2 bins, and given in Tab. 5.3.

$p(K)$ [GeV/c]	$\mathcal{A}_{\text{det}}$ [%]
$0 < p(K) < 10$	-1.37 ± 0.11
$10 < p(K) < 17.5$	-1.2 ± 0.10
$17.5 < p(K) < 22.5$	-1.15 ± 0.11
$22.5 < p(K) < 30$	-1.09 ± 0.12
$30 < p(K) < 50$	-0.88 ± 0.16
$50 < p(K) < 70$	-0.71 ± 0.29
$70 < p(K) < 100$	-0.33 ± 0.30
$100 < p(K) < 150$	0.18 ± 0.45

q^2 bin [GeV$^2/c^4$]	$\mathcal{A}_{\text{det}}$ [%]
$0.1 < q^2 < 0.98$	-0.010
$1.1 < q^2 < 2.5$	-0.011
$2.5 < q^2 < 4.0$	-0.011
$4.0 < q^2 < 6.0$	-0.011
$6.0 < q^2 < 8.0$	-0.011
$11.0 < q^2 < 12.5$	-0.011
$15.0 < q^2 < 17.0$	-0.012
$17.0 < q^2 < 19.0$	-0.012
$1.1 < q^2 < 6.0$	-0.011
$15.0 < q^2 < 19.0$	-0.012

Table 5.3: (Left) kaon detection asymmetry, depending on kaon momentum. (Right) resulting $K^+\pi^-$ detection asymmetry $\mathcal{A}_{\text{det}}$ for the different q^2 bins.

5.1.7 Summary of the systematic uncertainties

An overview of the systematic uncertainties for all the angular observables S_i in all the bins of q^2 is given in Tables 5.4, 5.5, 5.6, 5.7 and 5.8. The different systematic sources and their contribution to the total systematic uncertainty are shown. The total systematic uncertainty is calculated as quadratic sum of the individual contributions. The statistical uncertainty from a fit of the data, evaluated using `HESSE`, is given as well for comparison. The Tables 5.9, 5.10, 5.11, 5.12 and 5.13 provide the corresponding systematic uncertainties for the $P_i^{(\prime)}$ basis. Tables 5.14, 5.15, 5.16, 5.17 and 5.18 give the systematic uncertainties for the CP asymmetries A_i.

For the CP-average observable F_{L} the highest systematic uncertainty is the π-reweighting in almost every q^2-bins, for the other observables the peaking background systematic uncertainty is the biggest in almost each q^2-bins, while systematic uncertainties related to $m_{K\pi}$ follow right after. For the CP asymmetries and P_i' observables the systematic uncertainty on the peaking backgrounds is again the main contribution to the total systematic uncertainty.

$1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0307	0.0375	0.0497	0.0457	0.0294	0.0460	0.0500	0.0405
π reweighting	0.0139	0.0004	0.0001	0.0011	0.0023	0.0002	0.0001	0.0001
K reweighting	0.0000	0.0000	0.0013	0.0000	0.0001	0.0000	0.0000	0.0000
$p_\mathrm{T}(B^0)$ reweighting	0.0032	0.0000	0.0002	0.0001	0.0005	0.0000	0.0000	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0005	0.0001	0.0003	0.0003	0.0000	0.0000	0.0000	0.0000
N_tracks reweighting	0.0000	0.0002	0.0011	0.0003	0.0004	0.0000	0.0000	0.0000
higher order acc.	0.0001	0.0000	0.0006	0.0001	0.0001	0.0014	0.0011	0.0007
$\epsilon(q^2)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
peaking bkg.	0.0063	0.0032	0.0012	0.0041	0.0058	0.0055	0.0066	0.0037
angular bkg. model	0.0021	0.0010	0.0017	0.0005	0.0013	0.0002	0.0001	0.0004
sig. mass	0.0017	0.0001	0.0003	0.0005	0.0001	0.0003	0.0002	0.0001
$m_{K\pi}$ isobar	0.0003	0.0000	0.0001	0.0002	0.0000	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0042	0.0001	0.0013	0.0026	0.0004	0.0006	0.0001	0.0000
$m_{K\pi}$ eff.	0.0001	0.0017	0.0027	0.0002	0.0018	0.0007	0.0041	0.0021
acc. stat.	0.0012	0.0011	0.0016	0.0018	0.0007	0.0001	0.0001	0.0000
$\mathcal{A}_\mathrm{det}$	0.0000	0.0008	0.0001	0.0005	0.0002	0.0005	0.0005	0.0004
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0001	0.0000	0.0001	0.0000	0.0001	0.0001	0.0000
$\sigma_\mathrm{syst.}$	0.0165	0.0040	0.0044	0.0054	0.0067	0.0058	0.0079	0.0043

$15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0267	0.0335	0.0364	0.0349	0.0279	0.0414	0.0425	0.0402
π reweighting	0.0040	0.0021	0.0009	0.0015	0.0017	0.0000	0.0000	0.0000
K reweighting	0.0003	0.0004	0.0000	0.0001	0.0003	0.0000	0.0000	0.0000
$p_\mathrm{T}(B^0)$ reweighting	0.0029	0.0002	0.0003	0.0004	0.0015	0.0000	0.0000	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0011	0.0004	0.0003	0.0003	0.0003	0.0000	0.0000	0.0000
N_tracks reweighting	0.0003	0.0002	0.0003	0.0002	0.0005	0.0000	0.0000	0.0000
higher order acc.	0.0003	0.0010	0.0002	0.0002	0.0004	0.0050	0.0003	0.0002
$\epsilon(q^2)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
peaking bkg.	0.0046	0.0050	0.0037	0.0032	0.0056	0.0025	0.0014	0.0012
angular bkg. model	0.0005	0.0008	0.0015	0.0010	0.0000	0.0001	0.0000	0.0007
sig. mass	0.0017	0.0037	0.0005	0.0019	0.0027	0.0000	0.0000	0.0001
$m_{K\pi}$ isobar	0.0003	0.0011	0.0009	0.0013	0.0011	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0013	0.0042	0.0035	0.0048	0.0042	0.0000	0.0000	0.0000
$m_{K\pi}$ eff.	0.0016	0.0034	0.0036	0.0060	0.0036	0.0031	0.0026	0.0008
acc. stat.	0.0029	0.0039	0.0023	0.0031	0.0022	0.0001	0.0001	0.0000
$\mathcal{A}_\mathrm{det}$	0.0000	0.0004	0.0009	0.0004	0.0013	0.0005	0.0003	0.0007
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0000	0.0001	0.0000	0.0001	0.0000	0.0000	0.0001
$\sigma_\mathrm{syst.}$	0.0079	0.0095	0.0071	0.0094	0.0090	0.0065	0.0030	0.0018

Table 5.4: Summary of systematic uncertainties for the *CP*-averaged observables S_i in the q^2 bins $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ and $15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

$0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0436	0.0608	0.0658	0.0569	0.0563	0.0580	0.0738	0.0576
π reweighting	0.0139	0.0010	0.0005	0.0030	0.0003	0.0003	0.0000	0.0001
K reweighting	0.0035	0.0010	0.0003	0.0010	0.0008	0.0002	0.0001	0.0001
$p_\mathrm{T}(B^0)$ reweighting	0.0009	0.0002	0.0003	0.0004	0.0003	0.0000	0.0001	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0019	0.0001	0.0019	0.0004	0.0019	0.0002	0.0000	0.0001
N_tracks reweighting	0.0010	0.0000	0.0005	0.0003	0.0022	0.0001	0.0002	0.0000
higher order acc.	0.0037	0.0007	0.0042	0.0162	0.0004	0.0036	0.0003	0.0017
$\epsilon(q^2)$	0.0025	0.0014	0.0037	0.0014	0.0028	0.0001	0.0003	0.0000
peaking bkg.	0.0064	0.0023	0.0039	0.0040	0.0062	0.0038	0.0066	0.0030
angular bkg. model	0.0003	0.0010	0.0007	0.0002	0.0001	0.0001	0.0000	0.0006
sig. mass	0.0009	0.0001	0.0000	0.0008	0.0005	0.0000	0.0000	0.0000
$m_{K\pi}$ isobar	0.0002	0.0000	0.0001	0.0004	0.0001	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0004	0.0000	0.0003	0.0009	0.0003	0.0001	0.0000	0.0000
$m_{K\pi}$ eff.	0.0007	0.0008	0.0025	0.0011	0.0005	0.0019	0.0033	0.0009
acc. stat.	0.0029	0.0038	0.0040	0.0045	0.0038	0.0003	0.0001	0.0000
$\mathcal{A}_\mathrm{det}$	0.0000	0.0001	0.0007	0.0000	0.0013	0.0008	0.0003	0.0003
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0000	0.0001	0.0000	0.0001	0.0001	0.0000	0.0000
$\sigma_\mathrm{syst.}$	0.0168	0.0051	0.0086	0.0177	0.0085	0.0057	0.0074	0.0036

$1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0679	0.0744	0.0939	0.0872	0.0596	0.0883	0.0977	0.0741
π reweighting	0.0149	0.0002	0.0017	0.0006	0.0077	0.0001	0.0001	0.0001
K reweighting	0.0006	0.0002	0.0006	0.0004	0.0004	0.0000	0.0000	0.0000
$p_\mathrm{T}(B^0)$ reweighting	0.0028	0.0000	0.0002	0.0005	0.0014	0.0000	0.0000	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0004	0.0001	0.0001	0.0006	0.0003	0.0000	0.0000	0.0000
N_tracks reweighting	0.0002	0.0002	0.0010	0.0003	0.0000	0.0000	0.0000	0.0000
higher order acc.	0.0108	0.0015	0.0007	0.0065	0.0020	0.0030	0.0030	0.0013
$\epsilon(q^2)$	0.0088	0.0005	0.0029	0.0005	0.0043	0.0001	0.0001	0.0000
peaking bkg.	0.0035	0.0029	0.0012	0.0040	0.0075	0.0011	0.0034	0.0042
angular bkg. model	0.0033	0.0031	0.0004	0.0010	0.0009	0.0005	0.0004	0.0013
sig. mass	0.0021	0.0001	0.0001	0.0006	0.0011	0.0005	0.0001	0.0001
$m_{K\pi}$ isobar	0.0004	0.0000	0.0000	0.0002	0.0001	0.0001	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0024	0.0000	0.0002	0.0011	0.0009	0.0003	0.0001	0.0000
$m_{K\pi}$ eff.	0.0022	0.0006	0.0003	0.0029	0.0006	0.0009	0.0008	0.0003
acc. stat.	0.0018	0.0015	0.0025	0.0026	0.0013	0.0002	0.0001	0.0001
$\mathcal{A}_\mathrm{det}$	0.0000	0.0004	0.0025	0.0012	0.0004	0.0009	0.0005	0.0000
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0001	0.0003	0.0001	0.0000	0.0001	0.0001	0.0000
$\sigma_\mathrm{syst.}$	0.0216	0.0049	0.0052	0.0089	0.0121	0.0036	0.0047	0.0047

Table 5.5: Summary of systematic uncertainties for the *CP*-averaged observables S_i in the q^2 bins $0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$ and $1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$.

$2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0857	0.0694	0.1162	0.0952	0.0661	0.1017	0.1124	0.0847
π reweighting	0.0118	0.0005	0.0007	0.0017	0.0043	0.0003	0.0002	0.0002
K reweighting	0.0001	0.0001	0.0010	0.0000	0.0001	0.0000	0.0000	0.0000
$p_\mathrm{T}(B^0)$ reweighting	0.0027	0.0001	0.0005	0.0001	0.0009	0.0001	0.0001	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0005	0.0001	0.0003	0.0004	0.0001	0.0000	0.0000	0.0000
N_tracks reweighting	0.0003	0.0002	0.0010	0.0001	0.0005	0.0000	0.0000	0.0000
higher order acc.	0.0022	0.0003	0.0017	0.0025	0.0000	0.0005	0.0009	0.0012
$\epsilon(q^2)$	0.0089	0.0002	0.0002	0.0019	0.0034	0.0002	0.0002	0.0001
peaking bkg.	0.0024	0.0048	0.0052	0.0051	0.0033	0.0050	0.0048	0.0060
angular bkg. model	0.0013	0.0024	0.0003	0.0010	0.0014	0.0002	0.0002	0.0019
sig. mass	0.0022	0.0003	0.0003	0.0007	0.0003	0.0001	0.0000	0.0002
$m_{K\pi}$ isobar	0.0004	0.0000	0.0001	0.0002	0.0000	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0050	0.0002	0.0011	0.0024	0.0006	0.0005	0.0001	0.0001
$m_{K\pi}$ eff.	0.0010	0.0034	0.0023	0.0025	0.0031	0.0020	0.0033	0.0025
acc. stat.	0.0013	0.0012	0.0018	0.0022	0.0009	0.0001	0.0001	0.0000
$\mathcal{A}_\mathrm{det}$	0.0000	0.0012	0.0001	0.0001	0.0002	0.0003	0.0008	0.0024
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003
$\sigma_\mathrm{syst.}$	0.0165	0.0066	0.0065	0.0076	0.0074	0.0055	0.0059	0.0073

$4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0513	0.0646	0.0802	0.0747	0.0493	0.0785	0.0878	0.0674
π reweighting	0.0126	0.0011	0.0021	0.0020	0.0020	0.0002	0.0002	0.0002
K reweighting	0.0004	0.0001	0.0010	0.0004	0.0001	0.0000	0.0000	0.0000
$p_\mathrm{T}(B^0)$ reweighting	0.0032	0.0001	0.0004	0.0005	0.0005	0.0000	0.0000	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0004	0.0001	0.0002	0.0000	0.0001	0.0000	0.0000	0.0000
N_tracks reweighting	0.0004	0.0000	0.0008	0.0001	0.0006	0.0000	0.0000	0.0000
higher order acc.	0.0033	0.0003	0.0015	0.0020	0.0011	0.0001	0.0006	0.0008
$\epsilon(q^2)$	0.0089	0.0005	0.0009	0.0020	0.0013	0.0002	0.0000	0.0001
peaking bkg.	0.0041	0.0050	0.0069	0.0085	0.0013	0.0037	0.0032	0.0014
angular bkg. model	0.0010	0.0022	0.0010	0.0010	0.0005	0.0002	0.0002	0.0033
sig. mass	0.0013	0.0004	0.0004	0.0013	0.0007	0.0001	0.0001	0.0004
$m_{K\pi}$ isobar	0.0002	0.0000	0.0001	0.0004	0.0000	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0035	0.0003	0.0020	0.0047	0.0006	0.0004	0.0001	0.0000
$m_{K\pi}$ eff.	0.0033	0.0028	0.0002	0.0012	0.0012	0.0023	0.0031	0.0016
acc. stat.	0.0012	0.0012	0.0015	0.0018	0.0008	0.0001	0.0000	0.0000
$\mathcal{A}_\mathrm{det}$	0.0000	0.0018	0.0018	0.0006	0.0002	0.0004	0.0000	0.0007
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0002	0.0002	0.0001	0.0000	0.0001	0.0000	0.0001
$\sigma_\mathrm{syst.}$	0.0174	0.0067	0.0082	0.0107	0.0035	0.0044	0.0045	0.0041

Table 5.6: Summary of systematic uncertainties for the *CP*-averaged observables S_i in the q^2 bins $2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$ and $4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$.

$6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0442	0.0551	0.0606	0.0574	0.0390	0.0647	0.0682	0.0571
π reweighting	0.0130	0.0014	0.0020	0.0010	0.0058	0.0001	0.0001	0.0001
K reweighting	0.0008	0.0002	0.0008	0.0004	0.0001	0.0000	0.0000	0.0000
$p_\mathrm{T}(B^0)$ reweighting	0.0039	0.0001	0.0001	0.0005	0.0016	0.0000	0.0000	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0004	0.0000	0.0001	0.0001	0.0002	0.0000	0.0000	0.0000
N_tracks reweighting	0.0007	0.0001	0.0006	0.0001	0.0006	0.0000	0.0000	0.0000
higher order acc.	0.0034	0.0010	0.0014	0.0021	0.0021	0.0009	0.0013	0.0020
$\epsilon(q^2)$	0.0043	0.0003	0.0001	0.0004	0.0020	0.0001	0.0001	0.0001
peaking bkg.	0.0023	0.0100	0.0097	0.0082	0.0040	0.0009	0.0057	0.0016
angular bkg. model	0.0012	0.0039	0.0016	0.0034	0.0008	0.0003	0.0001	0.0018
sig. mass	0.0006	0.0006	0.0006	0.0019	0.0011	0.0002	0.0002	0.0002
$m_{K\pi}$ isobar	0.0001	0.0000	0.0002	0.0004	0.0001	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0016	0.0005	0.0019	0.0047	0.0014	0.0003	0.0001	0.0000
$m_{K\pi}$ eff.	0.0031	0.0016	0.0030	0.0054	0.0019	0.0023	0.0018	0.0034
acc. stat.	0.0013	0.0013	0.0012	0.0018	0.0009	0.0001	0.0000	0.0000
$\mathcal{A}_\mathrm{det}$	0.0000	0.0007	0.0004	0.0014	0.0005	0.0004	0.0005	0.0012
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0001	0.0000	0.0002	0.0001	0.0000	0.0001	0.0001
$\sigma_\mathrm{syst.}$	0.0153	0.0111	0.0109	0.0120	0.0083	0.0027	0.0061	0.0048

$11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0420	0.0434	0.0676	0.0613	0.0363	0.0689	0.0636	0.0581
π reweighting	0.0108	0.0015	0.0002	0.0016	0.0066	0.0000	0.0000	0.0000
K reweighting	0.0012	0.0001	0.0002	0.0002	0.0008	0.0000	0.0000	0.0000
$p_\mathrm{T}(B^0)$ reweighting	0.0046	0.0005	0.0006	0.0009	0.0028	0.0000	0.0000	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0010	0.0000	0.0001	0.0001	0.0006	0.0000	0.0000	0.0000
N_tracks reweighting	0.0003	0.0007	0.0003	0.0005	0.0002	0.0000	0.0000	0.0000
higher order acc.	0.0032	0.0005	0.0003	0.0028	0.0018	0.0014	0.0023	0.0008
$\epsilon(q^2)$	0.0013	0.0009	0.0005	0.0005	0.0009	0.0001	0.0002	0.0000
peaking bkg.	0.0018	0.0031	0.0080	0.0058	0.0022	0.0047	0.0042	0.0051
angular bkg. model	0.0007	0.0011	0.0014	0.0024	0.0021	0.0003	0.0002	0.0034
sig. mass	0.0011	0.0014	0.0005	0.0019	0.0019	0.0000	0.0000	0.0000
$m_{K\pi}$ isobar	0.0001	0.0001	0.0002	0.0005	0.0003	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0006	0.0012	0.0021	0.0044	0.0027	0.0000	0.0000	0.0000
$m_{K\pi}$ eff.	0.0006	0.0013	0.0013	0.0023	0.0005	0.0017	0.0020	0.0019
acc. stat.	0.0017	0.0017	0.0012	0.0019	0.0012	0.0000	0.0000	0.0000
$\mathcal{A}_\mathrm{det}$	0.0000	0.0015	0.0011	0.0003	0.0003	0.0001	0.0002	0.0006
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0001
$\sigma_\mathrm{syst.}$	0.0126	0.0050	0.0087	0.0092	0.0089	0.0052	0.0052	0.0065

Table 5.7: Summary of systematic uncertainties for the *CP*-averaged observables S_i in the q^2 bins $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$ and $11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$.

$15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0371	0.0402	0.0475	0.0469	0.0351	0.0547	0.0552	0.0501
π reweighting	0.0059	0.0020	0.0006	0.0018	0.0031	0.0000	0.0000	0.0000
K reweighting	0.0008	0.0002	0.0000	0.0000	0.0005	0.0000	0.0000	0.0000
$p_\mathrm{T}(B^0)$ reweighting	0.0032	0.0001	0.0001	0.0005	0.0017	0.0000	0.0000	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0010	0.0002	0.0003	0.0004	0.0004	0.0000	0.0000	0.0000
N_tracks reweighting	0.0010	0.0001	0.0002	0.0004	0.0006	0.0000	0.0000	0.0000
higher order acc.	0.0018	0.0013	0.0009	0.0022	0.0004	0.0016	0.0018	0.0024
$\epsilon(q^2)$	0.0034	0.0033	0.0030	0.0024	0.0014	0.0002	0.0002	0.0001
peaking bkg.	0.0034	0.0018	0.0049	0.0068	0.0054	0.0050	0.0025	0.0034
angular bkg. model	0.0011	0.0002	0.0012	0.0005	0.0002	0.0000	0.0000	0.0004
sig. mass	0.0013	0.0039	0.0003	0.0019	0.0023	0.0000	0.0001	0.0001
$m_{K\pi}$ isobar	0.0003	0.0007	0.0006	0.0010	0.0008	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0006	0.0013	0.0012	0.0018	0.0015	0.0000	0.0000	0.0000
$m_{K\pi}$ eff.	0.0020	0.0017	0.0009	0.0012	0.0007	0.0010	0.0010	0.0013
acc. stat.	0.0025	0.0030	0.0021	0.0028	0.0020	0.0000	0.0000	0.0000
$\mathcal{A}_\mathrm{det}$	0.0000	0.0004	0.0008	0.0009	0.0010	0.0012	0.0006	0.0011
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
$\sigma_\mathrm{syst.}$	0.0093	0.0071	0.0066	0.0089	0.0076	0.0055	0.0033	0.0045

$17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$								
σ	F_L	S_3	S_4	S_5	A_FB	S_7	S_8	S_9
$\sigma_\mathrm{stat.}$	0.0453	0.0642	0.0538	0.0527	0.0445	0.0682	0.0656	0.0576
π reweighting	0.0021	0.0019	0.0010	0.0011	0.0005	0.0000	0.0000	0.0000
K reweighting	0.0002	0.0006	0.0001	0.0002	0.0001	0.0000	0.0000	0.0000
$p_\mathrm{T}(B^0)$ reweighting	0.0024	0.0005	0.0005	0.0003	0.0011	0.0000	0.0000	0.0000
$\chi^2_\mathrm{Vtx.}$ reweighting	0.0013	0.0007	0.0004	0.0002	0.0002	0.0000	0.0000	0.0000
N_tracks reweighting	0.0015	0.0001	0.0009	0.0001	0.0016	0.0000	0.0000	0.0000
higher order acc.	0.0084	0.0025	0.0002	0.0031	0.0059	0.0112	0.0027	0.0015
$\epsilon(q^2)$	0.0226	0.0102	0.0062	0.0044	0.0063	0.0000	0.0000	0.0000
peaking bkg.	0.0037	0.0049	0.0061	0.0038	0.0074	0.0037	0.0043	0.0033
angular bkg. model	0.0003	0.0059	0.0002	0.0001	0.0002	0.0001	0.0002	0.0018
sig. mass	0.0022	0.0072	0.0011	0.0020	0.0034	0.0003	0.0000	0.0001
$m_{K\pi}$ isobar	0.0010	0.0035	0.0024	0.0026	0.0024	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0004	0.0015	0.0010	0.0011	0.0010	0.0000	0.0000	0.0000
$m_{K\pi}$ eff.	0.0008	0.0015	0.0014	0.0009	0.0013	0.0056	0.0007	0.0020
acc. stat.	0.0044	0.0067	0.0037	0.0046	0.0037	0.0000	0.0000	0.0000
$\mathcal{A}_\mathrm{det}$	0.0000	0.0007	0.0009	0.0001	0.0016	0.0006	0.0003	0.0005
$\mathcal{A}_\mathrm{prod}$	0.0000	0.0001	0.0001	0.0000	0.0002	0.0001	0.0000	0.0001
$\sigma_\mathrm{syst.}$	0.0252	0.0170	0.0102	0.0089	0.0130	0.0131	0.0051	0.0045

Table 5.8: Summary of systematic uncertainties for the *CP*-averaged observables S_i in the q^2 bins $15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$ and $17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

	$1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$						
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	0.2418	0.0643	0.1307	0.1074	0.0987	0.0997	0.1081
π reweighting	0.0016	0.0008	0.0003	0.0040	0.0034	0.0009	0.0002
K reweighting	0.0004	0.0002	0.0001	0.0029	0.0001	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0004	0.0002	0.0001	0.0004	0.0014	0.0002	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0010	0.0001	0.0000	0.0005	0.0004	0.0001	0.0000
N_{tracks} reweighting	0.0014	0.0010	0.0000	0.0025	0.0006	0.0000	0.0000
higher order acc.	0.0007	0.0000	0.0026	0.0013	0.0000	0.0031	0.0030
$\epsilon(q^2)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
peaking bkg.	0.0426	0.0076	0.0149	0.0076	0.0104	0.0199	0.0084
angular bkg. model	0.0022	0.0052	0.0019	0.0034	0.0002	0.0005	0.0005
sig. mass	0.0019	0.0020	0.0002	0.0010	0.0021	0.0006	0.0001
$m_{K\pi}$ isobar	0.0001	0.0002	0.0000	0.0003	0.0005	0.0001	0.0000
$m_{K\pi}$ bkg.	0.0021	0.0027	0.0002	0.0037	0.0065	0.0016	0.0002
$m_{K\pi}$ eff.	0.0029	0.0007	0.0061	0.0035	0.0003	0.0035	0.0019
acc. stat.	0.0081	0.0018	0.0001	0.0036	0.0044	0.0002	0.0001
$\mathcal{A}_{\mathrm{det}}$	0.0049	0.0003	0.0011	0.0003	0.0010	0.0010	0.0011
$\mathcal{A}_{\mathrm{prod}}$	0.0006	0.0000	0.0001	0.0000	0.0001	0.0001	0.0001
$\sigma_{\mathrm{syst.}}$	0.0440	0.0100	0.0165	0.0119	0.0138	0.0205	0.0092

	$15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$						
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	0.1006	0.0264	0.0611	0.0769	0.0743	0.0872	0.0896
π reweighting	0.0031	0.0008	0.0000	0.0001	0.0012	0.0000	0.0000
K reweighting	0.0016	0.0001	0.0000	0.0002	0.0004	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0031	0.0003	0.0000	0.0018	0.0006	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0003	0.0004	0.0000	0.0002	0.0001	0.0000	0.0000
N_{tracks} reweighting	0.0008	0.0004	0.0000	0.0005	0.0005	0.0000	0.0000
higher order acc.	0.0036	0.0002	0.0002	0.0008	0.0003	0.0099	0.0005
$\epsilon(q^2)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
peaking bkg.	0.0120	0.0077	0.0047	0.0035	0.0093	0.0047	0.0075
angular bkg. model	0.0016	0.0004	0.0015	0.0035	0.0021	0.0002	0.0001
sig. mass	0.0094	0.0022	0.0003	0.0018	0.0046	0.0000	0.0003
$m_{K\pi}$ isobar	0.0031	0.0009	0.0000	0.0022	0.0030	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0134	0.0040	0.0000	0.0092	0.0127	0.0001	0.0000
$m_{K\pi}$ eff.	0.0120	0.0033	0.0013	0.0081	0.0086	0.0028	0.0018
acc. stat.	0.0111	0.0013	0.0001	0.0049	0.0066	0.0002	0.0002
$\mathcal{A}_{\mathrm{det}}$	0.0013	0.0010	0.0011	0.0020	0.0009	0.0010	0.0006
$\mathcal{A}_{\mathrm{prod}}$	0.0001	0.0001	0.0001	0.0002	0.0001	0.0001	0.0001
$\sigma_{\mathrm{syst.}}$	0.0270	0.0098	0.0052	0.0147	0.0201	0.0114	0.0077

Table 5.9: Summary of systematic uncertainties for the $P_i^{(\prime)}$ in the q^2 bins $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ and $15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

$0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$							
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	0.1649	0.0510	0.0780	0.1498	0.1302	0.1317	0.1673
π reweighting	0.0027	0.0017	0.0001	0.0055	0.0036	0.0003	0.0001
K reweighting	0.0028	0.0009	0.0000	0.0011	0.0002	0.0003	0.0002
$p_{\mathrm{T}}(B^0)$ reweighting	0.0005	0.0004	0.0000	0.0009	0.0001	0.0000	0.0001
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0004	0.0015	0.0000	0.0014	0.0002	0.0004	0.0001
N_{tracks} reweighting	0.0000	0.0017	0.0000	0.0019	0.0001	0.0001	0.0002
higher order acc.	0.0025	0.0009	0.0021	0.0118	0.0435	0.0079	0.0007
$\epsilon(q^2)$	0.0038	0.0024	0.0000	0.0093	0.0031	0.0004	0.0007
peaking bkg.	0.0057	0.0041	0.0044	0.0127	0.0249	0.0114	0.0052
angular bkg. model	0.0028	0.0000	0.0006	0.0018	0.0008	0.0002	0.0000
sig. mass	0.0007	0.0000	0.0002	0.0005	0.0029	0.0002	0.0001
$m_{K\pi}$ isobar	0.0000	0.0001	0.0000	0.0004	0.0012	0.0001	0.0000
$m_{K\pi}$ bkg.	0.0000	0.0001	0.0000	0.0005	0.0015	0.0001	0.0000
$m_{K\pi}$ eff.	0.0037	0.0022	0.0043	0.0026	0.0021	0.0042	0.0041
acc. stat.	0.0101	0.0035	0.0001	0.0097	0.0119	0.0008	0.0002
$\mathcal{A}_{\mathrm{det}}$	0.0002	0.0009	0.0004	0.0016	0.0000	0.0018	0.0007
$\mathcal{A}_{\mathrm{prod}}$	0.0000	0.0001	0.0001	0.0002	0.0000	0.0002	0.0001
$\sigma_{\mathrm{syst.}}$	0.0139	0.0070	0.0065	0.0231	0.0519	0.0147	0.0067

$1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$							
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	0.4388	0.1330	0.2237	0.1982	0.1858	0.1891	0.2071
π reweighting	0.0014	0.0035	0.0003	0.0032	0.0047	0.0009	0.0001
K reweighting	0.0012	0.0001	0.0001	0.0015	0.0010	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0000	0.0008	0.0000	0.0005	0.0004	0.0002	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0004	0.0001	0.0001	0.0003	0.0014	0.0000	0.0000
N_{tracks} reweighting	0.0013	0.0003	0.0001	0.0023	0.0007	0.0000	0.0000
higher order acc.	0.0103	0.0104	0.0046	0.0027	0.0170	0.0054	0.0069
$\epsilon(q^2)$	0.0032	0.0026	0.0001	0.0052	0.0029	0.0008	0.0003
peaking bkg.	0.0224	0.0214	0.0118	0.0175	0.0100	0.0093	0.0175
angular bkg. model	0.0233	0.0062	0.0042	0.0008	0.0013	0.0010	0.0004
sig. mass	0.0013	0.0059	0.0003	0.0005	0.0022	0.0005	0.0004
$m_{K\pi}$ isobar	0.0000	0.0009	0.0000	0.0001	0.0005	0.0002	0.0000
$m_{K\pi}$ bkg.	0.0004	0.0060	0.0000	0.0006	0.0032	0.0010	0.0002
$m_{K\pi}$ eff.	0.0143	0.0056	0.0071	0.0034	0.0074	0.0039	0.0143
acc. stat.	0.0098	0.0020	0.0002	0.0054	0.0060	0.0004	0.0002
$\mathcal{A}_{\mathrm{det}}$	0.0026	0.0006	0.0001	0.0052	0.0026	0.0019	0.0010
$\mathcal{A}_{\mathrm{prod}}$	0.0003	0.0001	0.0000	0.0006	0.0003	0.0002	0.0001
$\sigma_{\mathrm{syst.}}$	0.0384	0.0270	0.0151	0.0207	0.0232	0.0117	0.0237

Table 5.10: Summary of systematic uncertainties for the $P_i^{(\prime)}$ in the q^2 bins $0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$ and $1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$.

	$2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$						
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	1.1432	0.5744	0.8633	0.4373	0.2902	0.3215	0.3429
π reweighting	0.0031	0.0013	0.0006	0.0047	0.0042	0.0013	0.0001
K reweighting	0.0007	0.0004	0.0001	0.0026	0.0000	0.0001	0.0001
$p_{\mathrm{T}}(B^0)$ reweighting	0.0004	0.0004	0.0001	0.0001	0.0021	0.0003	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0012	0.0003	0.0000	0.0006	0.0007	0.0001	0.0000
N_{tracks} reweighting	0.0023	0.0012	0.0001	0.0027	0.0006	0.0000	0.0000
higher order acc.	0.0022	0.0028	0.0056	0.0030	0.0077	0.0015	0.0021
$\epsilon(q^2)$	0.0001	0.0006	0.0004	0.0044	0.0013	0.0010	0.0005
peaking bkg.	0.0009	0.0103	0.0186	0.0214	0.0151	0.0111	0.0248
angular bkg. model	0.0248	0.0027	0.0090	0.0003	0.0022	0.0004	0.0003
sig. mass	0.0017	0.0046	0.0017	0.0022	0.0037	0.0009	0.0004
$m_{K\pi}$ isobar	0.0002	0.0006	0.0000	0.0004	0.0007	0.0002	0.0001
$m_{K\pi}$ bkg.	0.0023	0.0081	0.0004	0.0051	0.0091	0.0022	0.0007
$m_{K\pi}$ eff.	0.0292	0.0042	0.0086	0.0021	0.0083	0.0056	0.0029
acc. stat.	0.0117	0.0023	0.0002	0.0046	0.0054	0.0004	0.0002
$\mathcal{A}_{\mathrm{det}}$	0.0192	0.0009	0.0197	0.0002	0.0002	0.0011	0.0023
$\mathcal{A}_{\mathrm{prod}}$	0.0022	0.0001	0.0023	0.0000	0.0000	0.0001	0.0003
$\sigma_{\mathrm{syst.}}$	0.0448	0.0154	0.0305	0.0241	0.0226	0.0129	0.0252

	$4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$						
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	0.3327	0.0844	0.1737	0.1642	0.1528	0.1613	0.1808
π reweighting	0.0060	0.0040	0.0004	0.0021	0.0073	0.0006	0.0000
K reweighting	0.0006	0.0006	0.0000	0.0025	0.0006	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0001	0.0011	0.0001	0.0007	0.0017	0.0002	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0008	0.0000	0.0000	0.0003	0.0004	0.0000	0.0000
N_{tracks} reweighting	0.0002	0.0013	0.0000	0.0016	0.0002	0.0000	0.0000
higher order acc.	0.0025	0.0051	0.0031	0.0051	0.0074	0.0001	0.0013
$\epsilon(q^2)$	0.0020	0.0030	0.0003	0.0023	0.0032	0.0005	0.0004
peaking bkg.	0.0108	0.0017	0.0104	0.0152	0.0136	0.0068	0.0093
angular bkg. model	0.0184	0.0005	0.0124	0.0014	0.0011	0.0006	0.0008
sig. mass	0.0028	0.0002	0.0009	0.0016	0.0037	0.0001	0.0002
$m_{K\pi}$ isobar	0.0003	0.0003	0.0000	0.0005	0.0011	0.0001	0.0000
$m_{K\pi}$ bkg.	0.0033	0.0040	0.0001	0.0060	0.0137	0.0011	0.0004
$m_{K\pi}$ eff.	0.0068	0.0057	0.0164	0.0072	0.0035	0.0024	0.0011
acc. stat.	0.0098	0.0020	0.0001	0.0031	0.0044	0.0002	0.0001
$\mathcal{A}_{\mathrm{det}}$	0.0094	0.0003	0.0017	0.0036	0.0013	0.0009	0.0001
$\mathcal{A}_{\mathrm{prod}}$	0.0011	0.0000	0.0002	0.0004	0.0001	0.0001	0.0000
$\sigma_{\mathrm{syst.}}$	0.0274	0.0105	0.0233	0.0198	0.0233	0.0074	0.0095

Table 5.11: Summary of systematic uncertainties for the $P_i^{(\prime)}$ in the q^2 bins $2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$ and $4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$.

$6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$							
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	0.2626	0.0593	0.1362	0.1227	0.1164	0.1312	0.1383
π reweighting	0.0053	0.0034	0.0004	0.0002	0.0052	0.0003	0.0001
K reweighting	0.0013	0.0006	0.0000	0.0019	0.0005	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0000	0.0011	0.0001	0.0009	0.0010	0.0001	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0002	0.0000	0.0000	0.0002	0.0004	0.0000	0.0000
N_{tracks} reweighting	0.0001	0.0004	0.0000	0.0010	0.0002	0.0000	0.0000
higher order acc.	0.0048	0.0003	0.0056	0.0018	0.0063	0.0018	0.0027
$\epsilon(q^2)$	0.0009	0.0009	0.0001	0.0011	0.0015	0.0002	0.0002
peaking bkg.	0.0132	0.0094	0.0083	0.0074	0.0155	0.0108	0.0045
angular bkg. model	0.0164	0.0001	0.0061	0.0022	0.0058	0.0006	0.0003
sig. mass	0.0029	0.0031	0.0004	0.0010	0.0041	0.0006	0.0000
$m_{K\pi}$ isobar	0.0002	0.0004	0.0000	0.0004	0.0010	0.0001	0.0000
$m_{K\pi}$ bkg.	0.0029	0.0048	0.0002	0.0050	0.0121	0.0007	0.0004
$m_{K\pi}$ eff.	0.0061	0.0046	0.0056	0.0006	0.0083	0.0022	0.0037
acc. stat.	0.0073	0.0011	0.0001	0.0025	0.0038	0.0002	0.0001
$\mathcal{A}_{\mathrm{det}}$	0.0033	0.0006	0.0028	0.0008	0.0028	0.0008	0.0009
$\mathcal{A}_{\mathrm{prod}}$	0.0004	0.0001	0.0003	0.0001	0.0003	0.0001	0.0001
$\sigma_{\mathrm{syst.}}$	0.0248	0.0126	0.0133	0.0101	0.0244	0.0113	0.0065

$11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$							
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	0.2808	0.0384	0.1152	0.1271	0.1162	0.1385	0.1285
π reweighting	0.0012	0.0012	0.0000	0.0021	0.0005	0.0000	0.0000
K reweighting	0.0007	0.0001	0.0000	0.0003	0.0001	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0000	0.0005	0.0000	0.0003	0.0005	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0003	0.0001	0.0000	0.0003	0.0004	0.0000	0.0000
N_{tracks} reweighting	0.0025	0.0001	0.0000	0.0005	0.0009	0.0000	0.0000
higher order acc.	0.0005	0.0005	0.0013	0.0001	0.0047	0.0028	0.0047
$\epsilon(q^2)$	0.0035	0.0000	0.0001	0.0013	0.0015	0.0002	0.0002
peaking bkg.	0.0082	0.0014	0.0075	0.0127	0.0091	0.0055	0.0016
angular bkg. model	0.0040	0.0032	0.0058	0.0030	0.0049	0.0004	0.0005
sig. mass	0.0049	0.0013	0.0000	0.0007	0.0039	0.0001	0.0002
$m_{K\pi}$ isobar	0.0005	0.0003	0.0000	0.0005	0.0011	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0033	0.0022	0.0001	0.0036	0.0077	0.0000	0.0000
$m_{K\pi}$ eff.	0.0051	0.0018	0.0026	0.0013	0.0028	0.0025	0.0008
acc. stat.	0.0057	0.0005	0.0000	0.0024	0.0038	0.0000	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0059	0.0003	0.0013	0.0023	0.0006	0.0002	0.0003
$\mathcal{A}_{\mathrm{prod}}$	0.0006	0.0000	0.0001	0.0002	0.0001	0.0000	0.0000
$\sigma_{\mathrm{syst.}}$	0.0152	0.0049	0.0100	0.0143	0.0151	0.0067	0.0050

Table 5.12: Summary of systematic uncertainties for the $P_i^{(\prime)}$ in the q^2 bins $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$ and $11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$.

$15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$							
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	0.1196	0.0316	0.0768	0.0994	0.1012	0.1149	0.1158
π reweighting	0.0023	0.0008	0.0000	0.0010	0.0010	0.0000	0.0000
K reweighting	0.0013	0.0000	0.0000	0.0002	0.0003	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0018	0.0004	0.0000	0.0011	0.0005	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0000	0.0002	0.0000	0.0001	0.0003	0.0000	0.0000
N_{tracks} reweighting	0.0003	0.0001	0.0000	0.0001	0.0003	0.0000	0.0000
higher order acc.	0.0028	0.0008	0.0036	0.0013	0.0037	0.0034	0.0038
$\epsilon(q^2)$	0.0078	0.0008	0.0000	0.0046	0.0029	0.0002	0.0001
peaking bkg.	0.0031	0.0042	0.0039	0.0123	0.0128	0.0042	0.0018
angular bkg. model	0.0004	0.0005	0.0011	0.0033	0.0021	0.0000	0.0000
sig. mass	0.0113	0.0015	0.0005	0.0012	0.0044	0.0000	0.0002
$m_{K\pi}$ isobar	0.0020	0.0006	0.0000	0.0014	0.0022	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0049	0.0015	0.0000	0.0036	0.0055	0.0000	0.0000
$m_{K\pi}$ eff.	0.0018	0.0016	0.0016	0.0056	0.0035	0.0025	0.0027
acc. stat.	0.0091	0.0010	0.0000	0.0043	0.0063	0.0000	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0012	0.0008	0.0017	0.0018	0.0019	0.0026	0.0012
$\mathcal{A}_{\mathrm{prod}}$	0.0001	0.0001	0.0002	0.0002	0.0002	0.0003	0.0001
$\sigma_{\mathrm{syst.}}$	0.0182	0.0054	0.0059	0.0160	0.0174	0.0065	0.0051

$17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$							
σ	P_1	P_2	P_3	P_4	P_5	P_6	P_8
$\sigma_{\mathrm{stat.}}$	0.2002	0.0436	0.0890	0.1148	0.1127	0.1424	0.1371
π reweighting	0.0038	0.0007	0.0000	0.0011	0.0014	0.0000	0.0000
K reweighting	0.0016	0.0002	0.0000	0.0002	0.0004	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0040	0.0002	0.0000	0.0024	0.0005	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0007	0.0005	0.0000	0.0002	0.0001	0.0000	0.0000
N_{tracks} reweighting	0.0019	0.0008	0.0000	0.0011	0.0009	0.0000	0.0000
higher order acc.	0.0012	0.0014	0.0022	0.0039	0.0031	0.0238	0.0058
$\epsilon(q^2)$	0.0079	0.0056	0.0001	0.0016	0.0016	0.0002	0.0002
peaking bkg.	0.0104	0.0016	0.0063	0.0094	0.0084	0.0081	0.0036
angular bkg. model	0.0162	0.0004	0.0028	0.0003	0.0001	0.0001	0.0003
sig. mass	0.0193	0.0020	0.0002	0.0036	0.0054	0.0003	0.0001
$m_{K\pi}$ isobar	0.0095	0.0019	0.0000	0.0055	0.0059	0.0001	0.0000
$m_{K\pi}$ bkg.	0.0070	0.0014	0.0000	0.0041	0.0044	0.0000	0.0000
$m_{K\pi}$ eff.	0.0031	0.0007	0.0031	0.0039	0.0043	0.0006	0.0058
acc. stat.	0.0195	0.0025	0.0000	0.0080	0.0097	0.0001	0.0001
$\mathcal{A}_{\mathrm{det}}$	0.0021	0.0012	0.0008	0.0018	0.0002	0.0012	0.0006
$\mathcal{A}_{\mathrm{prod}}$	0.0002	0.0001	0.0001	0.0002	0.0000	0.0001	0.0001
$\sigma_{\mathrm{syst.}}$	0.0371	0.0074	0.0079	0.0161	0.0168	0.0252	0.0090

Table 5.13: Summary of systematic uncertainties for the $P_i^{(\prime)}$ in the q^2 bins $15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$ and $17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

$1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0375	0.0497	0.0457	0.0294	0.0460	0.0500	0.0405
π reweighting	0.0004	0.0001	0.0011	0.0031	0.0002	0.0001	0.0001
K reweighting	0.0000	0.0013	0.0000	0.0001	0.0000	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0000	0.0002	0.0001	0.0007	0.0000	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0001	0.0003	0.0003	0.0000	0.0000	0.0000	0.0000
N_{tracks} reweighting	0.0002	0.0011	0.0003	0.0005	0.0000	0.0000	0.0000
higher order acc.	0.0000	0.0006	0.0001	0.0001	0.0014	0.0011	0.0007
$\epsilon(q^2)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
peaking bkg.	0.0032	0.0012	0.0041	0.0078	0.0055	0.0066	0.0037
angular bkg. model	0.0010	0.0017	0.0005	0.0018	0.0002	0.0001	0.0004
sig. mass	0.0001	0.0003	0.0005	0.0001	0.0003	0.0002	0.0001
$m_{K\pi}$ isobar	0.0000	0.0001	0.0002	0.0000	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0001	0.0013	0.0026	0.0006	0.0006	0.0001	0.0000
$m_{K\pi}$ eff.	0.0017	0.0027	0.0002	0.0023	0.0007	0.0041	0.0021
acc. stat.	0.0011	0.0016	0.0018	0.0010	0.0001	0.0001	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0001	0.0016	0.0002	0.0011	0.0008	0.0003	0.0007
$\mathcal{A}_{\mathrm{prod}}$	0.0000	0.0002	0.0000	0.0001	0.0001	0.0000	0.0001
$\sigma_{\mathrm{syst.}}$	0.0039	0.0047	0.0054	0.0090	0.0058	0.0078	0.0044

$15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0335	0.0364	0.0349	0.0279	0.0414	0.0425	0.0402
π reweighting	0.0021	0.0009	0.0015	0.0023	0.0000	0.0000	0.0000
K reweighting	0.0004	0.0000	0.0001	0.0004	0.0000	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0002	0.0003	0.0004	0.0019	0.0000	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0004	0.0003	0.0003	0.0004	0.0000	0.0000	0.0000
N_{tracks} reweighting	0.0002	0.0003	0.0002	0.0007	0.0000	0.0000	0.0000
higher order acc.	0.0010	0.0002	0.0002	0.0006	0.0050	0.0003	0.0002
$\epsilon(q^2)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
peaking bkg.	0.0050	0.0037	0.0032	0.0074	0.0025	0.0014	0.0012
angular bkg. model	0.0008	0.0015	0.0010	0.0000	0.0001	0.0000	0.0007
sig. mass	0.0037	0.0005	0.0019	0.0036	0.0000	0.0000	0.0001
$m_{K\pi}$ isobar	0.0011	0.0009	0.0013	0.0015	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0042	0.0035	0.0048	0.0056	0.0000	0.0000	0.0000
$m_{K\pi}$ eff.	0.0034	0.0036	0.0060	0.0047	0.0031	0.0026	0.0008
acc. stat.	0.0039	0.0023	0.0031	0.0030	0.0001	0.0001	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0020	0.0034	0.0039	0.0057	0.0006	0.0003	0.0006
$\mathcal{A}_{\mathrm{prod}}$	0.0002	0.0003	0.0004	0.0006	0.0001	0.0000	0.0001
$\sigma_{\mathrm{syst.}}$	0.0097	0.0078	0.0102	0.0132	0.0065	0.0030	0.0018

Table 5.14: Summary of systematic uncertainties for the *CP* asymmetries A_i in the q^2 bins $1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$ and $15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

$0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0608	0.0658	0.0569	0.0563	0.0580	0.0738	0.0576
π reweighting	0.0010	0.0005	0.0030	0.0005	0.0003	0.0000	0.0001
K reweighting	0.0010	0.0003	0.0010	0.0011	0.0002	0.0001	0.0001
$p_{\mathrm{T}}(B^0)$ reweighting	0.0002	0.0003	0.0004	0.0004	0.0000	0.0001	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0001	0.0019	0.0004	0.0025	0.0002	0.0000	0.0001
N_{tracks} reweighting	0.0000	0.0005	0.0003	0.0029	0.0001	0.0002	0.0000
higher order acc.	0.0007	0.0042	0.0162	0.0005	0.0036	0.0003	0.0017
$\epsilon(q^2)$	0.0014	0.0037	0.0014	0.0037	0.0001	0.0003	0.0000
peaking bkg.	0.0023	0.0039	0.0040	0.0082	0.0038	0.0066	0.0030
angular bkg. model	0.0010	0.0007	0.0002	0.0001	0.0001	0.0000	0.0006
sig. mass	0.0001	0.0000	0.0008	0.0007	0.0000	0.0000	0.0000
$m_{K\pi}$ isobar	0.0000	0.0001	0.0004	0.0002	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0000	0.0003	0.0009	0.0004	0.0001	0.0000	0.0000
$m_{K\pi}$ eff.	0.0008	0.0025	0.0011	0.0007	0.0019	0.0033	0.0009
acc. stat.	0.0038	0.0040	0.0045	0.0051	0.0003	0.0001	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0004	0.0008	0.0018	0.0000	0.0002	0.0008	0.0009
$\mathcal{A}_{\mathrm{prod}}$	0.0000	0.0001	0.0002	0.0000	0.0000	0.0001	0.0001
$\sigma_{\mathrm{syst.}}$	0.0051	0.0086	0.0178	0.0112	0.0056	0.0074	0.0037

$1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0744	0.0939	0.0872	0.0596	0.0883	0.0977	0.0741
π reweighting	0.0002	0.0017	0.0006	0.0103	0.0001	0.0001	0.0001
K reweighting	0.0002	0.0006	0.0004	0.0005	0.0000	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0000	0.0002	0.0005	0.0018	0.0000	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0001	0.0001	0.0006	0.0004	0.0000	0.0000	0.0000
N_{tracks} reweighting	0.0002	0.0010	0.0003	0.0001	0.0000	0.0000	0.0000
higher order acc.	0.0015	0.0007	0.0065	0.0027	0.0030	0.0030	0.0013
$\epsilon(q^2)$	0.0005	0.0029	0.0005	0.0057	0.0001	0.0001	0.0000
peaking bkg.	0.0029	0.0012	0.0040	0.0100	0.0011	0.0034	0.0042
angular bkg. model	0.0031	0.0004	0.0010	0.0011	0.0005	0.0004	0.0013
sig. mass	0.0001	0.0001	0.0006	0.0014	0.0005	0.0001	0.0001
$m_{K\pi}$ isobar	0.0000	0.0000	0.0002	0.0002	0.0001	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0000	0.0002	0.0011	0.0012	0.0003	0.0001	0.0000
$m_{K\pi}$ eff.	0.0006	0.0003	0.0029	0.0008	0.0009	0.0008	0.0003
acc. stat.	0.0015	0.0025	0.0026	0.0017	0.0002	0.0001	0.0001
$\mathcal{A}_{\mathrm{det}}$	0.0008	0.0008	0.0014	0.0027	0.0023	0.0010	0.0013
$\mathcal{A}_{\mathrm{prod}}$	0.0001	0.0001	0.0002	0.0003	0.0003	0.0001	0.0001
$\sigma_{\mathrm{syst.}}$	0.0049	0.0046	0.0089	0.0163	0.0042	0.0048	0.0048

Table 5.15: Summary of systematic uncertainties for the *CP* asymmetries A_i in the q^2 bins $0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$ and $1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$.

$2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0694	0.1162	0.0952	0.0661	0.1017	0.1124	0.0847
π reweighting	0.0005	0.0007	0.0017	0.0057	0.0003	0.0002	0.0002
K reweighting	0.0001	0.0010	0.0000	0.0001	0.0000	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0001	0.0005	0.0001	0.0013	0.0001	0.0001	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0001	0.0003	0.0004	0.0001	0.0000	0.0000	0.0000
N_{tracks} reweighting	0.0002	0.0010	0.0001	0.0006	0.0000	0.0000	0.0000
higher order acc.	0.0003	0.0017	0.0025	0.0000	0.0005	0.0009	0.0012
$\epsilon(q^2)$	0.0002	0.0002	0.0019	0.0045	0.0002	0.0002	0.0001
peaking bkg.	0.0048	0.0052	0.0051	0.0044	0.0050	0.0048	0.0060
angular bkg. model	0.0024	0.0003	0.0010	0.0019	0.0002	0.0002	0.0019
sig. mass	0.0003	0.0003	0.0007	0.0003	0.0001	0.0000	0.0002
$m_{K\pi}$ isobar	0.0000	0.0001	0.0002	0.0001	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0002	0.0011	0.0024	0.0008	0.0005	0.0001	0.0001
$m_{K\pi}$ eff.	0.0034	0.0023	0.0025	0.0042	0.0020	0.0033	0.0025
acc. stat.	0.0012	0.0018	0.0022	0.0012	0.0001	0.0001	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0004	0.0025	0.0002	0.0017	0.0007	0.0003	0.0010
$\mathcal{A}_{\mathrm{prod}}$	0.0000	0.0003	0.0000	0.0002	0.0001	0.0000	0.0001
$\sigma_{\mathrm{syst.}}$	0.0065	0.0070	0.0076	0.0100	0.0055	0.0059	0.0070

$4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0646	0.0802	0.0747	0.0493	0.0785	0.0878	0.0674
π reweighting	0.0011	0.0021	0.0020	0.0027	0.0002	0.0002	0.0002
K reweighting	0.0001	0.0010	0.0004	0.0002	0.0000	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0001	0.0004	0.0005	0.0006	0.0000	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0001	0.0002	0.0000	0.0002	0.0000	0.0000	0.0000
N_{tracks} reweighting	0.0000	0.0008	0.0001	0.0008	0.0000	0.0000	0.0000
higher order acc.	0.0003	0.0015	0.0020	0.0015	0.0001	0.0006	0.0008
$\epsilon(q^2)$	0.0005	0.0009	0.0020	0.0017	0.0002	0.0000	0.0001
peaking bkg.	0.0050	0.0069	0.0085	0.0017	0.0037	0.0032	0.0014
angular bkg. model	0.0022	0.0010	0.0010	0.0006	0.0002	0.0002	0.0033
sig. mass	0.0004	0.0004	0.0013	0.0009	0.0001	0.0001	0.0004
$m_{K\pi}$ isobar	0.0000	0.0001	0.0004	0.0001	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0003	0.0020	0.0047	0.0008	0.0004	0.0001	0.0000
$m_{K\pi}$ eff.	0.0028	0.0002	0.0012	0.0015	0.0023	0.0031	0.0016
acc. stat.	0.0012	0.0015	0.0018	0.0011	0.0001	0.0000	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0004	0.0023	0.0015	0.0003	0.0002	0.0018	0.0003
$\mathcal{A}_{\mathrm{prod}}$	0.0000	0.0003	0.0002	0.0000	0.0000	0.0002	0.0000
$\sigma_{\mathrm{syst.}}$	0.0064	0.0083	0.0108	0.0047	0.0044	0.0048	0.0041

Table 5.16: Summary of systematic uncertainties for the *CP* asymmetries A_i in the q^2 bins $2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$ and $4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$.

$6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0551	0.0606	0.0574	0.0390	0.0647	0.0682	0.0571
π reweighting	0.0014	0.0020	0.0010	0.0077	0.0001	0.0001	0.0001
K reweighting	0.0002	0.0008	0.0004	0.0002	0.0000	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0001	0.0001	0.0005	0.0022	0.0000	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0000	0.0001	0.0001	0.0003	0.0000	0.0000	0.0000
N_{tracks} reweighting	0.0001	0.0006	0.0001	0.0008	0.0000	0.0000	0.0000
higher order acc.	0.0010	0.0014	0.0021	0.0028	0.0009	0.0013	0.0020
$\epsilon(q^2)$	0.0003	0.0001	0.0004	0.0027	0.0001	0.0001	0.0001
peaking bkg.	0.0100	0.0097	0.0082	0.0053	0.0009	0.0057	0.0016
angular bkg. model	0.0039	0.0016	0.0034	0.0010	0.0003	0.0001	0.0018
sig. mass	0.0006	0.0006	0.0019	0.0015	0.0002	0.0002	0.0002
$m_{K\pi}$ isobar	0.0000	0.0002	0.0004	0.0002	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0005	0.0019	0.0047	0.0018	0.0003	0.0001	0.0000
$m_{K\pi}$ eff.	0.0016	0.0030	0.0054	0.0025	0.0023	0.0018	0.0034
acc. stat.	0.0013	0.0012	0.0018	0.0012	0.0001	0.0000	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0005	0.0032	0.0027	0.0022	0.0005	0.0009	0.0003
$\mathcal{A}_{\mathrm{prod}}$	0.0001	0.0004	0.0003	0.0002	0.0001	0.0001	0.0000
$\sigma_{\mathrm{syst.}}$	0.0111	0.0113	0.0123	0.0113	0.0027	0.0062	0.0047

$11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0434	0.0676	0.0613	0.0363	0.0689	0.0636	0.0581
π reweighting	0.0015	0.0002	0.0016	0.0088	0.0000	0.0000	0.0000
K reweighting	0.0001	0.0002	0.0002	0.0010	0.0000	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0005	0.0006	0.0009	0.0038	0.0000	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0000	0.0001	0.0001	0.0009	0.0000	0.0000	0.0000
N_{tracks} reweighting	0.0007	0.0003	0.0005	0.0002	0.0000	0.0000	0.0000
higher order acc.	0.0005	0.0003	0.0028	0.0024	0.0014	0.0023	0.0008
$\epsilon(q^2)$	0.0009	0.0005	0.0005	0.0012	0.0001	0.0002	0.0000
peaking bkg.	0.0031	0.0080	0.0058	0.0030	0.0047	0.0042	0.0051
angular bkg. model	0.0011	0.0014	0.0024	0.0028	0.0003	0.0002	0.0034
sig. mass	0.0014	0.0005	0.0019	0.0025	0.0000	0.0000	0.0000
$m_{K\pi}$ isobar	0.0001	0.0002	0.0005	0.0004	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0012	0.0021	0.0044	0.0037	0.0000	0.0000	0.0000
$m_{K\pi}$ eff.	0.0013	0.0013	0.0023	0.0007	0.0017	0.0020	0.0019
acc. stat.	0.0017	0.0012	0.0019	0.0016	0.0000	0.0000	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0021	0.0032	0.0037	0.0048	0.0016	0.0001	0.0000
$\mathcal{A}_{\mathrm{prod}}$	0.0002	0.0003	0.0004	0.0005	0.0002	0.0000	0.0000
$\sigma_{\mathrm{syst.}}$	0.0052	0.0092	0.0099	0.0128	0.0055	0.0052	0.0065

Table 5.17: Summary of systematic uncertainties for the *CP* asymmetries A_i in the q^2 bins $6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$ and $11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$.

$15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0402	0.0475	0.0469	0.0351	0.0547	0.0552	0.0501
π reweighting	0.0020	0.0006	0.0018	0.0042	0.0000	0.0000	0.0000
K reweighting	0.0002	0.0000	0.0000	0.0007	0.0000	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0001	0.0001	0.0005	0.0023	0.0000	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0002	0.0003	0.0004	0.0006	0.0000	0.0000	0.0000
N_{tracks} reweighting	0.0001	0.0002	0.0004	0.0007	0.0000	0.0000	0.0000
higher order acc.	0.0013	0.0009	0.0022	0.0005	0.0016	0.0018	0.0024
$\epsilon(q^2)$	0.0033	0.0030	0.0024	0.0019	0.0002	0.0002	0.0001
peaking bkg.	0.0018	0.0049	0.0068	0.0072	0.0050	0.0025	0.0034
angular bkg. model	0.0002	0.0012	0.0005	0.0003	0.0000	0.0000	0.0004
sig. mass	0.0039	0.0003	0.0019	0.0031	0.0000	0.0001	0.0001
$m_{K\pi}$ isobar	0.0007	0.0006	0.0010	0.0010	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0013	0.0012	0.0018	0.0020	0.0000	0.0000	0.0000
$m_{K\pi}$ eff.	0.0017	0.0009	0.0012	0.0009	0.0010	0.0010	0.0013
acc. stat.	0.0030	0.0021	0.0028	0.0026	0.0000	0.0000	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0017	0.0038	0.0037	0.0065	0.0007	0.0000	0.0002
$\mathcal{A}_{\mathrm{prod}}$	0.0002	0.0004	0.0004	0.0007	0.0001	0.0000	0.0000
$\sigma_{\mathrm{syst.}}$	0.0072	0.0076	0.0096	0.0120	0.0054	0.0032	0.0043

$17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$							
σ	A_3	A_4	A_5	A_6	A_7	A_8	A_9
$\sigma_{\mathrm{stat.}}$	0.0642	0.0538	0.0527	0.0445	0.0682	0.0656	0.0576
π reweighting	0.0019	0.0010	0.0011	0.0006	0.0000	0.0000	0.0000
K reweighting	0.0006	0.0001	0.0002	0.0001	0.0000	0.0000	0.0000
$p_{\mathrm{T}}(B^0)$ reweighting	0.0005	0.0005	0.0003	0.0015	0.0000	0.0000	0.0000
$\chi^2_{\mathrm{Vtx.}}$ reweighting	0.0007	0.0004	0.0002	0.0002	0.0000	0.0000	0.0000
N_{tracks} reweighting	0.0001	0.0009	0.0001	0.0021	0.0000	0.0000	0.0000
higher order acc.	0.0025	0.0002	0.0031	0.0079	0.0112	0.0027	0.0015
$\epsilon(q^2)$	0.0102	0.0062	0.0044	0.0084	0.0000	0.0000	0.0000
peaking bkg.	0.0049	0.0061	0.0038	0.0098	0.0037	0.0043	0.0033
angular bkg. model	0.0059	0.0002	0.0001	0.0003	0.0001	0.0002	0.0018
sig. mass	0.0072	0.0011	0.0020	0.0045	0.0003	0.0000	0.0001
$m_{K\pi}$ isobar	0.0035	0.0024	0.0026	0.0032	0.0000	0.0000	0.0000
$m_{K\pi}$ bkg.	0.0015	0.0010	0.0011	0.0013	0.0000	0.0000	0.0000
$m_{K\pi}$ eff.	0.0015	0.0014	0.0009	0.0018	0.0056	0.0007	0.0020
acc. stat.	0.0067	0.0037	0.0046	0.0049	0.0000	0.0000	0.0000
$\mathcal{A}_{\mathrm{det}}$	0.0023	0.0032	0.0039	0.0050	0.0005	0.0002	0.0011
$\mathcal{A}_{\mathrm{prod}}$	0.0002	0.0003	0.0004	0.0005	0.0001	0.0000	0.0001
$\sigma_{\mathrm{syst.}}$	0.0171	0.0106	0.0097	0.0179	0.0131	0.0051	0.0046

Table 5.18: Summary of systematic uncertainties for the *CP* asymmetries A_i in the q^2 bins $15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$ and $17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$.

5.2 Results

5.2.1 The previous measurements of the $B^0 \to K^{*0}\mu^+\mu^-$ angular observables

The study of the $B^0 \to K^{*0}\mu^+\mu^-$ angular observables started with the B-factories, BaBar [6] and Belle [7], which published, in 2008 and 2009 respectively, the measurements of the $F_{\rm L}$ and $A_{\rm FB}$ observables. In 2011, the CDF collaboration published their measurement adding two new angular observables, $A_{\rm T}^2$ and $A_{\rm Im}$ [8]. In 2013, using the first collected 1fb^{-1} of data, the LHCb experiment entered the game first publishing the most precise measurements of the $F_{\rm L}$, $A_{\rm FB}$, S_3, S_9, A_9, $A_{\rm T}^2$ and $A_{\rm T}^{\rm Re}$ observables [9], and then measuring for the first time a set of less form factor independent observables, the P_i' [10]. All the measurements of the angular observables agreed with the Standard Model, except for the P_5', which showed a discrepancy of 3.7 standard deviations from the prediction in one q^2 bin. The collaborations ATLAS [11] and CMS [12] have also published their results on the $F_{\rm L}$ and $A_{\rm FB}$ observables. All these results, compared with the theoretical predictions, are shown in Figures. 5.5 and 5.6.

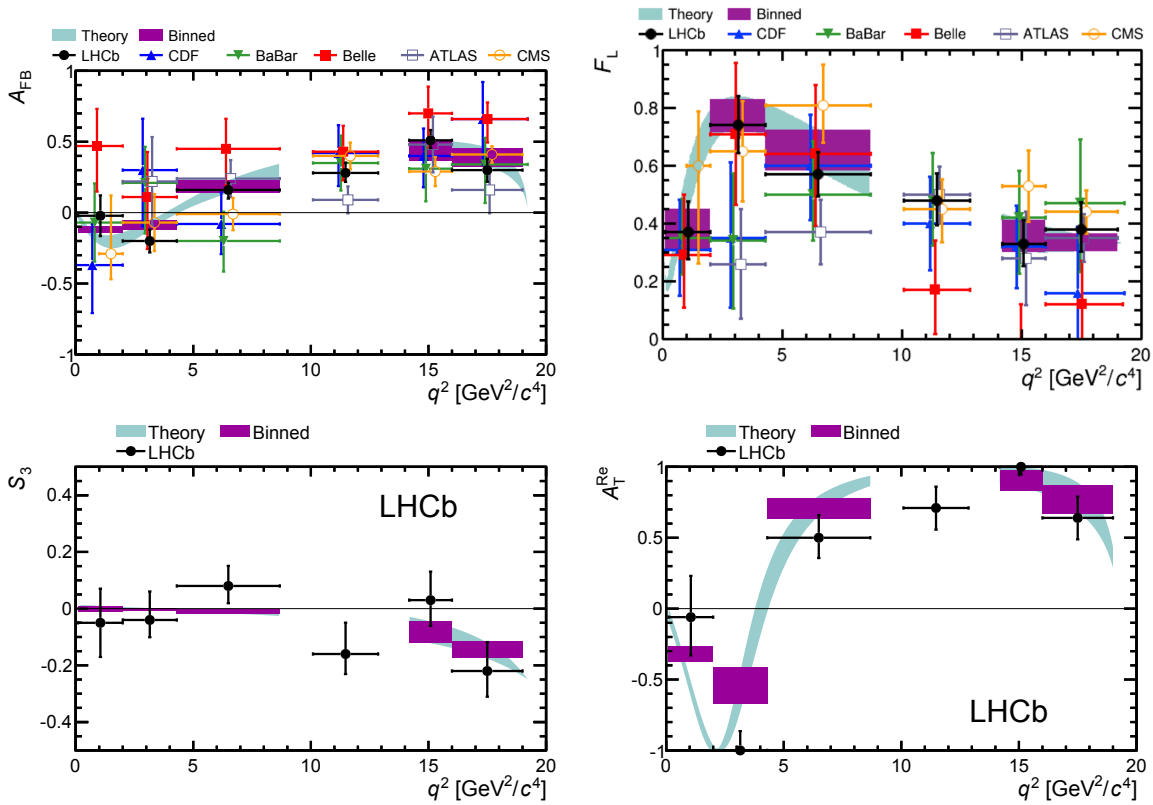


Figure 5.5: On the top, comparison of the measurements of the $A_{\rm FB}$ (left) and $F_{\rm L}$ (right) angular observables performed by different collaborations.The Standard Model predictions, in blue bands and red boxes, are explained in Ref. [13]. On the bottom, the S_3 and A_T^{Re} measurements from LHCb with $1fb^{-1}$ are shown.

Recently, using the data recorded at a centre-of-mass energy of $\sqrt{s} = 8$ TeV, the CMS collaboration has published an update of its measurement of the $F_{\rm L}$ and $A_{\rm FB}$ observables [15].

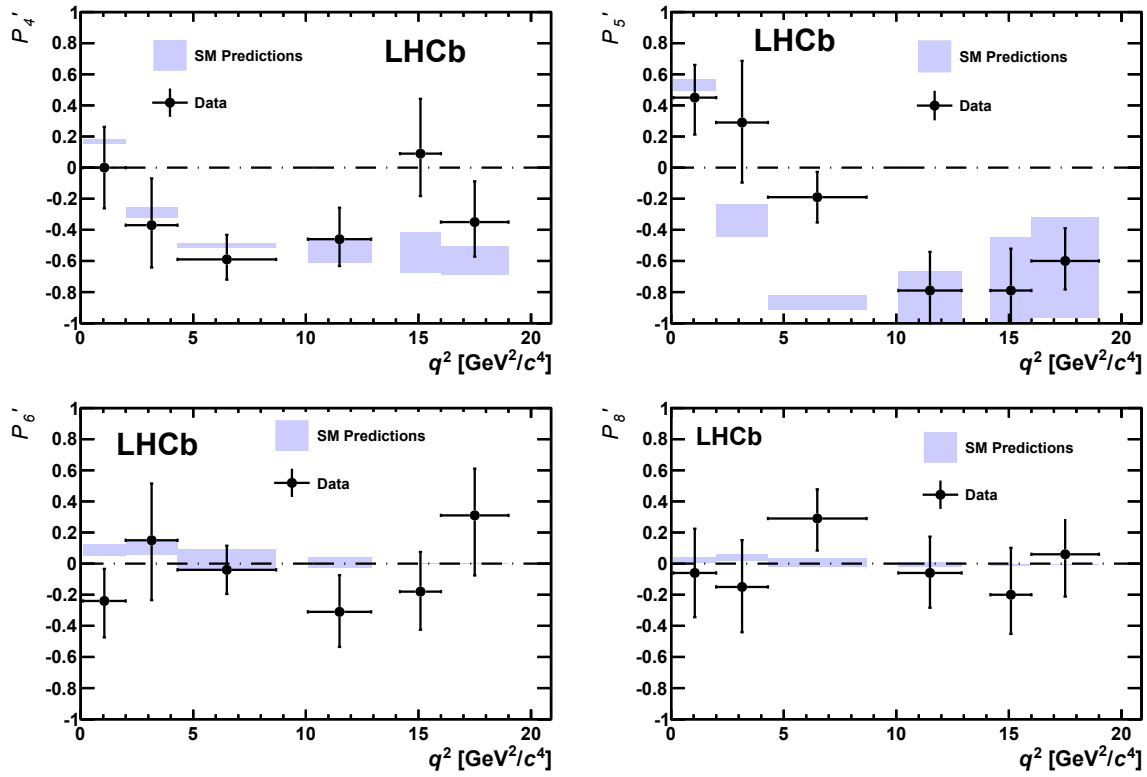


Figure 5.6: Measurements of the P_i' observables by LHCb with 1 fb^{-1} of data. The Standard Model predictions from [14] are shown in light blue boxes.

5.2.2 The new LHCb measurement based on 3 fb^{-1} of data

The angular analysis presented in this thesis makes use of the whole dataset collected by the LHCb collaboration during the first run of the LHC, corresponding to an integrated luminosity of 3 fb^{-1}. Due to a larger number of events available respect to the previous publication, the analysis has been performed on a larger number os smaller q^2 bins, giving the opportunity to have a better handling on the q^2 dependencies of the angular observables. Moreover, this larger number of events allows for the first time to extract the full set of angular observables without using folding techniques, and so it enables the extraction of the correlations between the different observables. These correlations are important for a correct treatment of the experimental results in a global fit.

The projections of the fitted probability density function on the decay angles, the reconstructed B^0 mass and $m_{K\pi}$ are given in Figures. 5.7 to 5.10 for different q^2 bins. The projections for the two large q^2 bins, $1.1 < q^2 < 6\,\mathrm{GeV}^2/c^4$ and $15 < q^2 < 19\,\mathrm{GeV}^2/c^4$ are given in Figures 5.11, where a weighted fit approach is used, since the acceptance significantly varies over these large bins.

The results of the fit for the CP-averaged observables in each q^2 bin are reported in Table 5.19 and in Figure 5.12. There is a good agreement with the Standard Model predictions for every observables. It should be noted, though, that the measured value of A_{FB} has the tendency to be lower than the prediction by about 1σ in the $1.1{<}q^2{<}6$ GeV$^2/c^4$ region.

A measurement of the CP-asymmetries observables, $A_{3,4,5,6,7,8,9}$, has also been made and the results, given in Table 5.20, are all compatible with Standard Model expectations.

Finally, the results for the less form factor dependent $P_i^{(\prime)}$ observables described in Section. 1.3.4 are given in Table. 5.21. The result of the P_5' observable is particularly interesting and shown in Figure. 5.13. The new analysis on 3 fb^{-1} of data confirms the deviation seen in the previous analysis [10] at the same level: indeed, a local discrepancy of the order of 2.9 standard deviations is observed in each of the $4 < q^2 < 6$ GeV$^2/c^4$ and $6 < q^2 < 8$ GeV$^2/c^4$ bins.

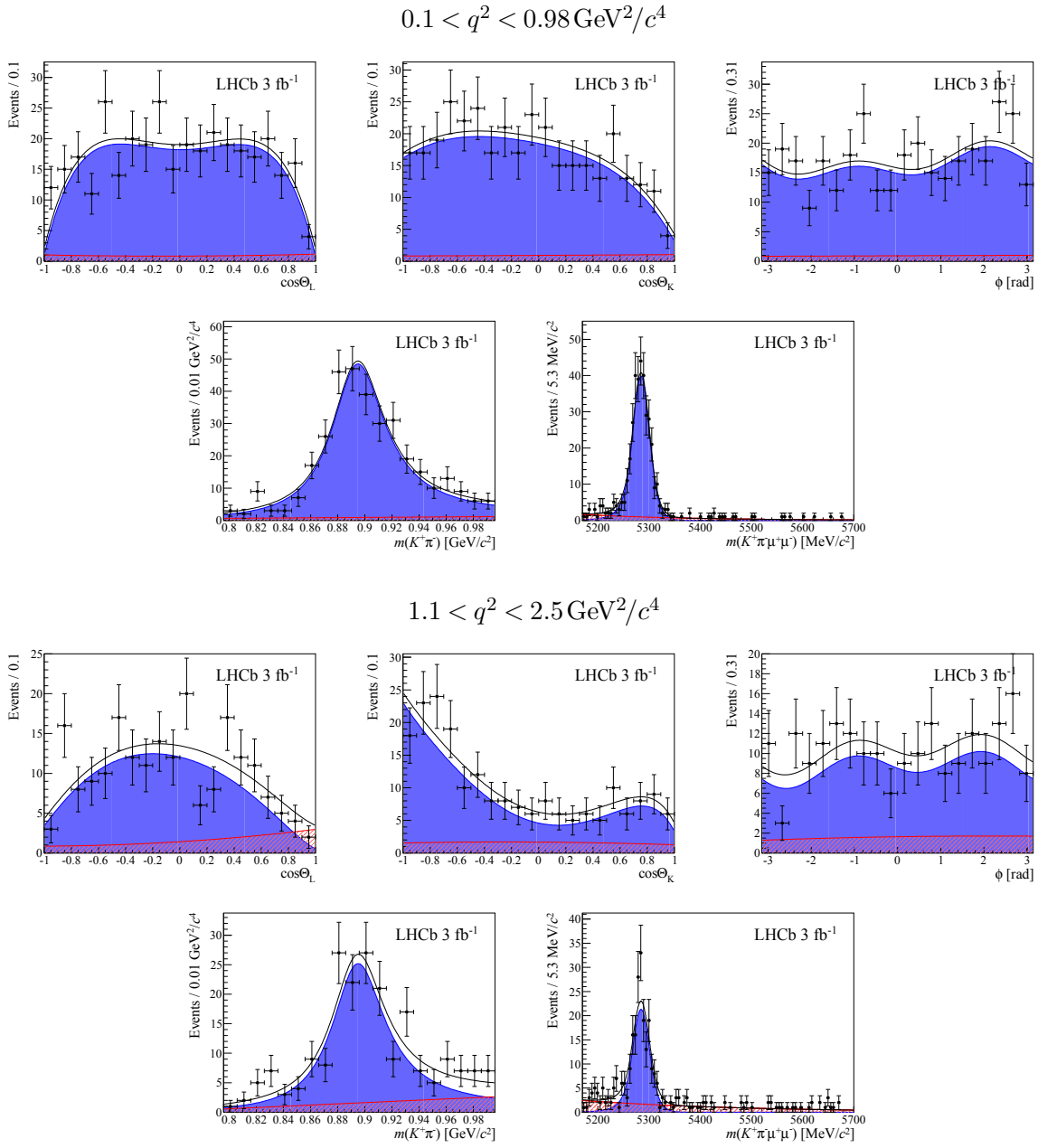


Figure 5.7: Projections of the fitted probability density function on the decay angles, $m_{K\pi}$ and the reconstructed B^0 mass in bins of q^2.

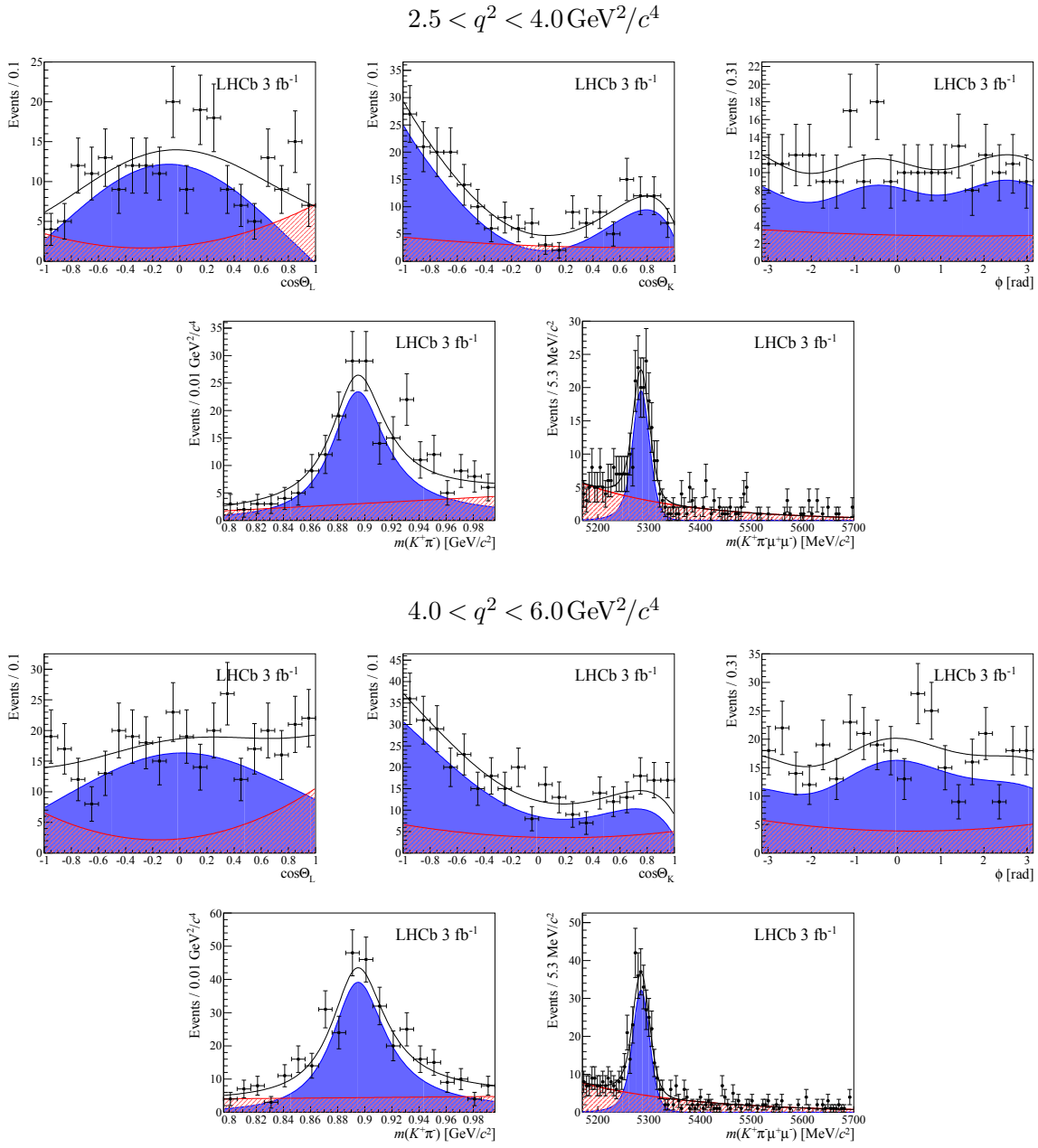


Figure 5.8: Projections of the fitted probability density function on the decay angles, $m_{K\pi}$ and the reconstructed B^0 mass in bins of q^2.

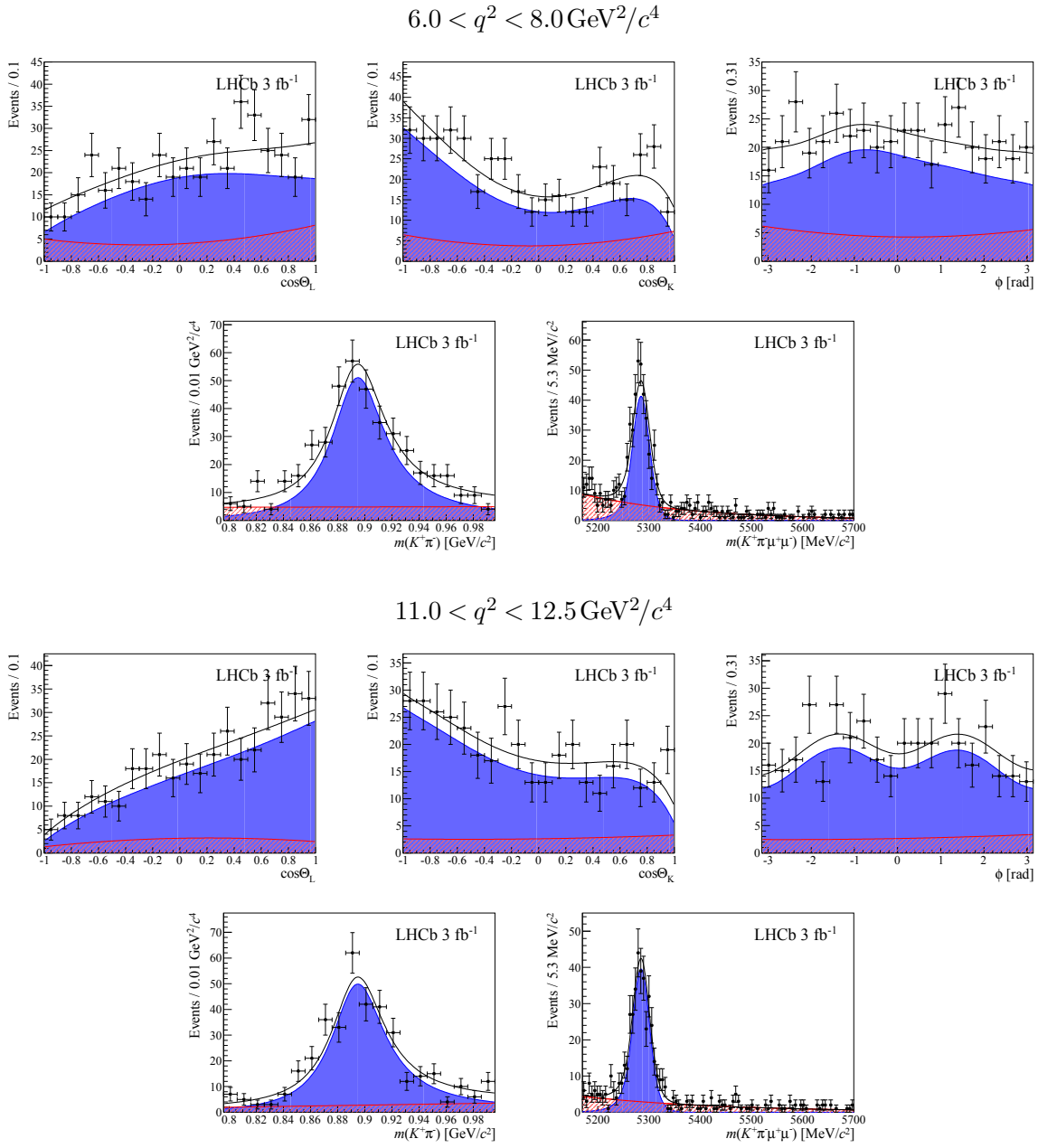


Figure 5.9: Projections of the fitted probability density function on the decay angles, $m_{K\pi}$ and the reconstructed B^0 mass in bins of q^2.

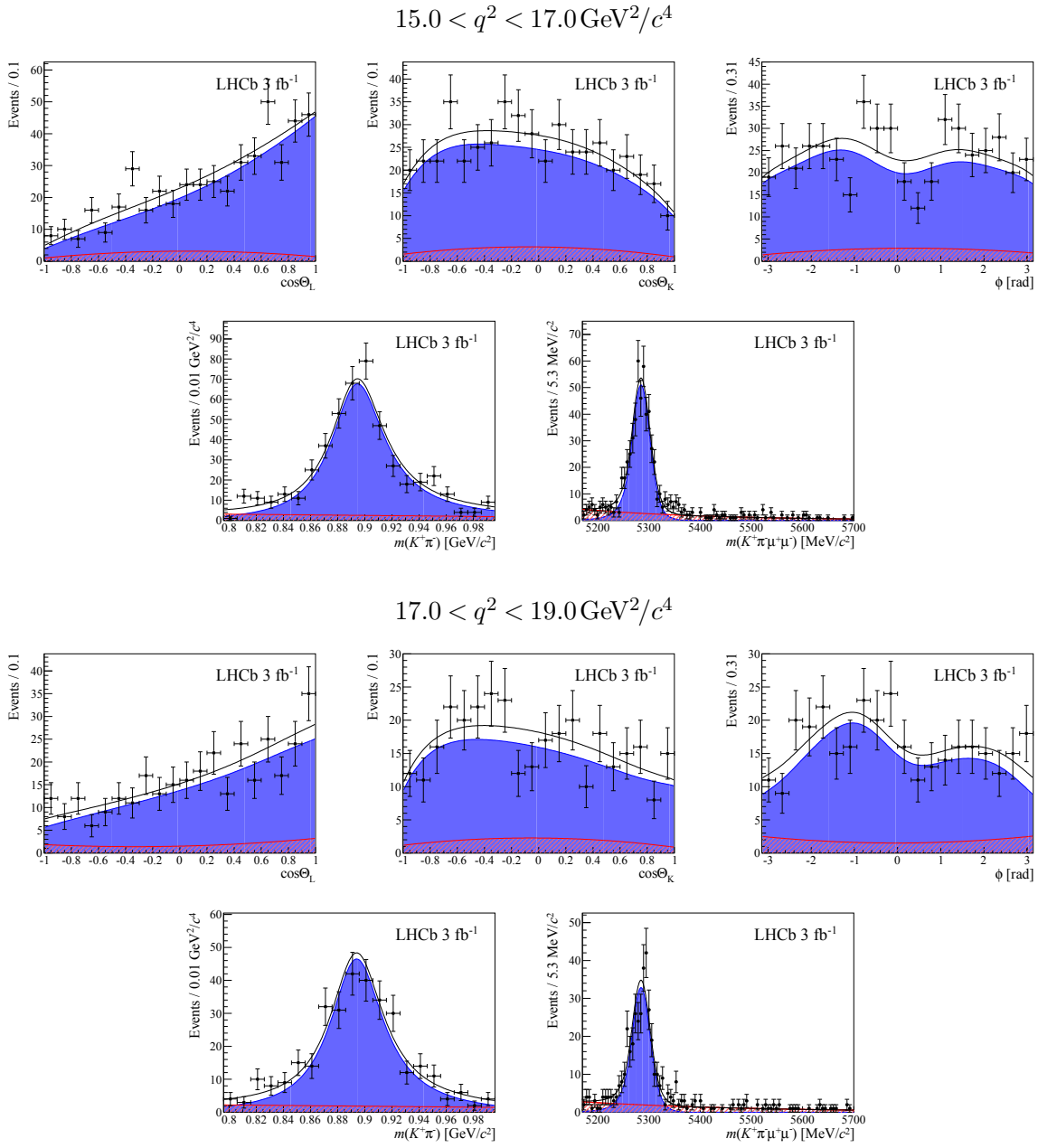


Figure 5.10: Projections of the fitted probability density function on the decay angles, $m_{K\pi}$ and the reconstructed B^0 mass in bins of q^2.

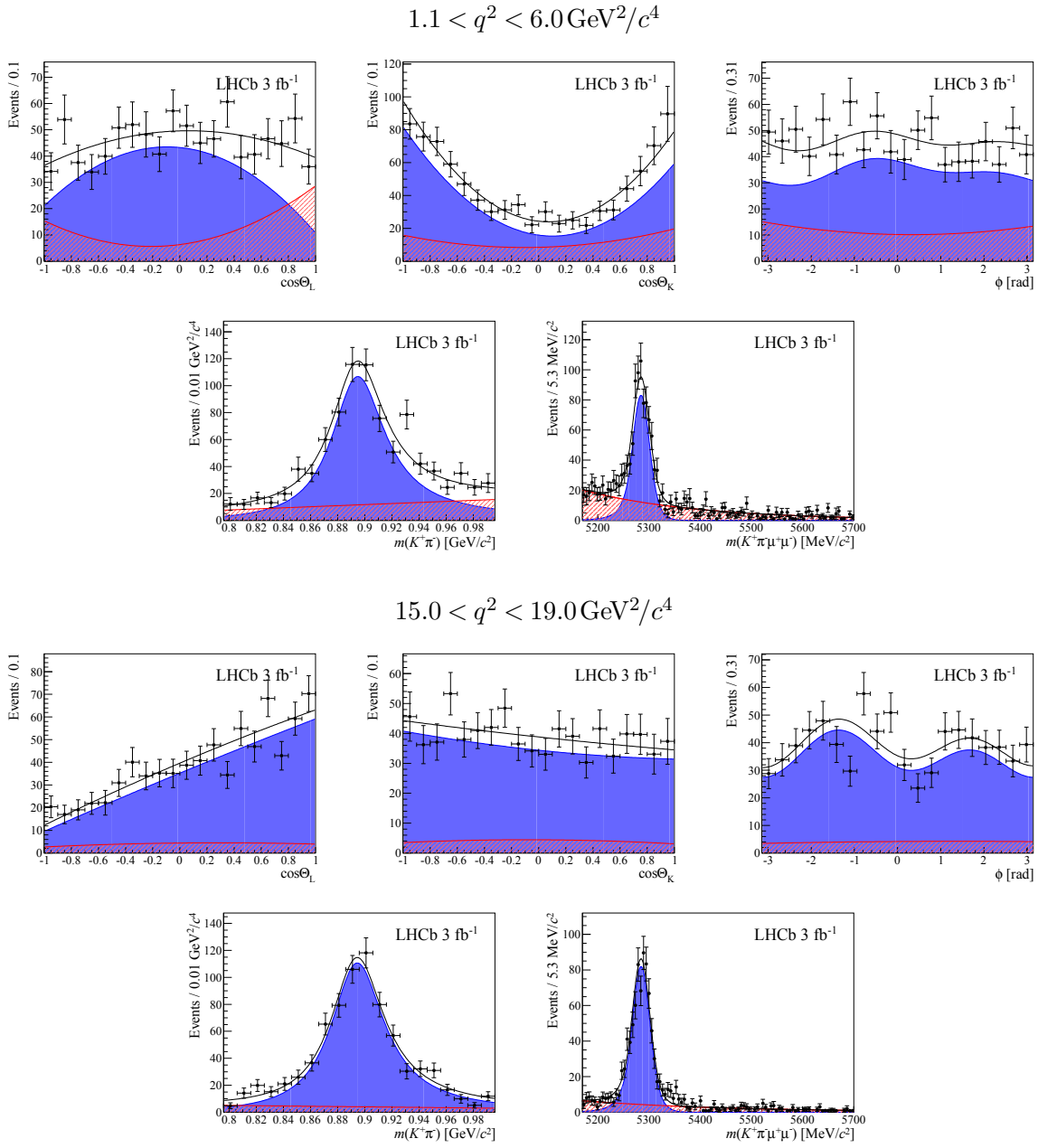


Figure 5.11: Projections of the fitted probability density function on the decay angles, $m_{K\pi}$ and the reconstructed B^0 mass in bins of q^2.

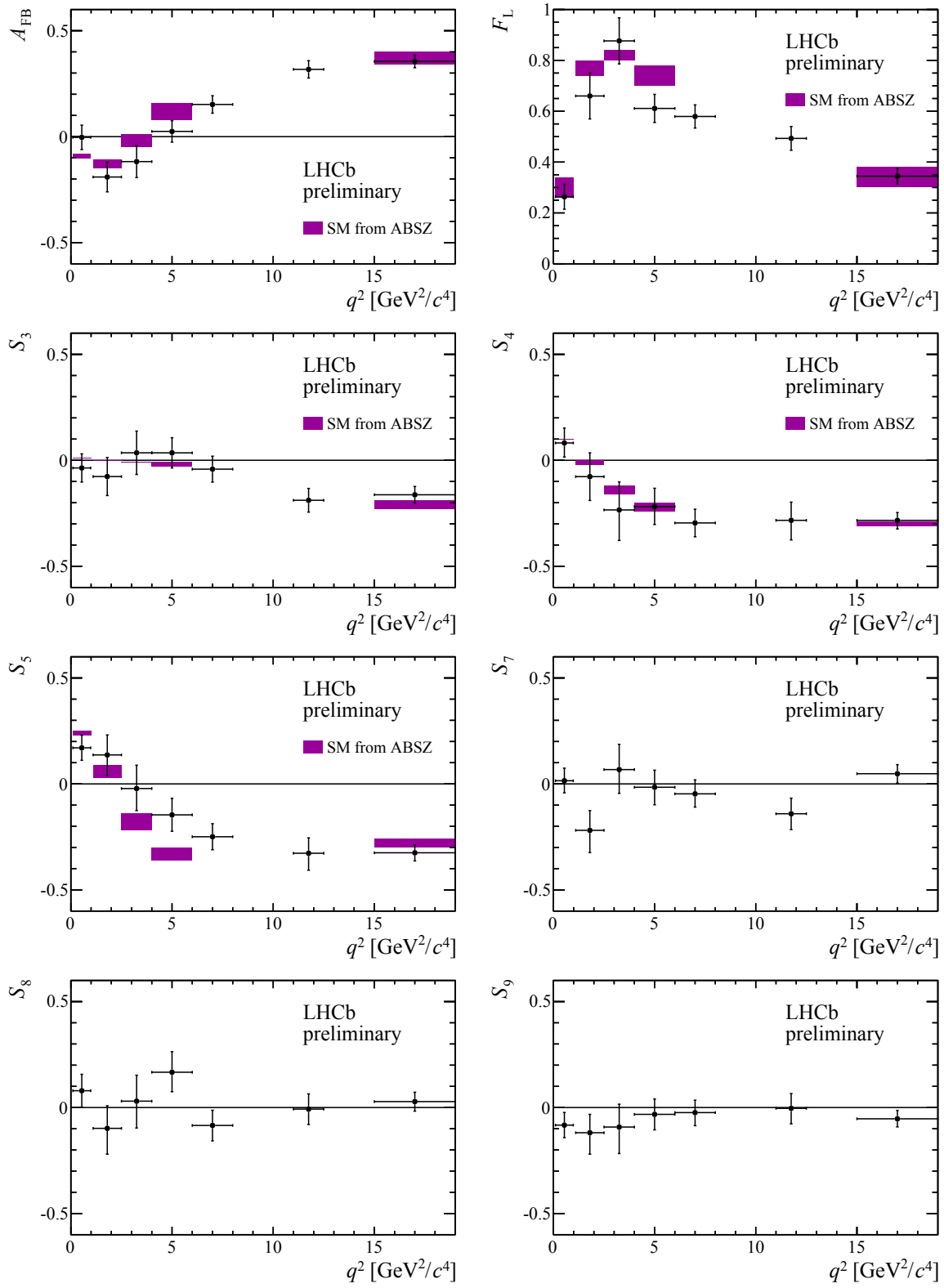


Figure 5.12: The CP-average observables in bins of q^2. The boxes are the SM prediction from Ref. [16].

$0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.263^{+0.045}_{-0.044} \pm 0.017$
S_3	$-0.036^{+0.063}_{-0.063} \pm 0.005$
S_4	$0.082^{+0.068}_{-0.069} \pm 0.009$
S_5	$0.170^{+0.059}_{-0.058} \pm 0.018$
A_FB	$-0.003^{+0.058}_{-0.057} \pm 0.009$
S_7	$0.015^{+0.059}_{-0.059} \pm 0.006$
S_8	$0.079^{+0.076}_{-0.075} \pm 0.007$
S_9	$-0.083^{+0.058}_{-0.057} \pm 0.004$

$1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.660^{+0.083}_{-0.077} \pm 0.022$
S_3	$-0.077^{+0.087}_{-0.105} \pm 0.005$
S_4	$-0.077^{+0.111}_{-0.113} \pm 0.005$
S_5	$0.137^{+0.099}_{-0.094} \pm 0.009$
A_FB	$-0.191^{+0.068}_{-0.080} \pm 0.012$
S_7	$-0.219^{+0.094}_{-0.104} \pm 0.004$
S_8	$-0.098^{+0.108}_{-0.123} \pm 0.005$
S_9	$-0.119^{+0.087}_{-0.104} \pm 0.005$

$2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.876^{+0.109}_{-0.097} \pm 0.017$
S_3	$0.035^{+0.098}_{-0.089} \pm 0.007$
S_4	$-0.234^{+0.127}_{-0.144} \pm 0.006$
S_5	$-0.022^{+0.110}_{-0.103} \pm 0.008$
A_FB	$-0.118^{+0.082}_{-0.090} \pm 0.007$
S_7	$0.068^{+0.120}_{-0.112} \pm 0.005$
S_8	$0.030^{+0.129}_{-0.131} \pm 0.006$
S_9	$-0.092^{+0.105}_{-0.125} \pm 0.007$

$4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.611^{+0.052}_{-0.053} \pm 0.017$
S_3	$0.035^{+0.069}_{-0.068} \pm 0.007$
S_4	$-0.219^{+0.086}_{-0.084} \pm 0.008$
S_5	$-0.146^{+0.077}_{-0.078} \pm 0.011$
A_FB	$0.025^{+0.051}_{-0.052} \pm 0.004$
S_7	$-0.016^{+0.081}_{-0.080} \pm 0.004$
S_8	$0.167^{+0.094}_{-0.091} \pm 0.004$
S_9	$-0.032^{+0.071}_{-0.071} \pm 0.004$

$6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.579^{+0.046}_{-0.046} \pm 0.015$
S_3	$-0.042^{+0.058}_{-0.059} \pm 0.011$
S_4	$-0.296^{+0.063}_{-0.067} \pm 0.011$
S_5	$-0.249^{+0.059}_{-0.060} \pm 0.012$
A_FB	$0.152^{+0.041}_{-0.040} \pm 0.008$
S_7	$-0.047^{+0.068}_{-0.066} \pm 0.003$
S_8	$-0.085^{+0.072}_{-0.070} \pm 0.006$
S_9	$-0.024^{+0.059}_{-0.060} \pm 0.005$

$11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.493^{+0.049}_{-0.047} \pm 0.013$
S_3	$-0.189^{+0.054}_{-0.058} \pm 0.005$
S_4	$-0.283^{+0.084}_{-0.095} \pm 0.009$
S_5	$-0.327^{+0.076}_{-0.079} \pm 0.009$
A_FB	$0.318^{+0.044}_{-0.040} \pm 0.009$
S_7	$-0.141^{+0.072}_{-0.074} \pm 0.005$
S_8	$-0.007^{+0.070}_{-0.072} \pm 0.005$
S_9	$-0.004^{+0.070}_{-0.073} \pm 0.006$

$15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.349^{+0.039}_{-0.039} \pm 0.009$
S_3	$-0.142^{+0.044}_{-0.049} \pm 0.007$
S_4	$-0.321^{+0.055}_{-0.074} \pm 0.007$
S_5	$-0.316^{+0.051}_{-0.057} \pm 0.009$
A_FB	$0.411^{+0.041}_{-0.037} \pm 0.008$
S_7	$0.061^{+0.058}_{-0.058} \pm 0.005$
S_8	$0.003^{+0.061}_{-0.061} \pm 0.003$
S_9	$-0.019^{+0.054}_{-0.056} \pm 0.004$

$17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.354^{+0.049}_{-0.048} \pm 0.025$
S_3	$-0.188^{+0.074}_{-0.084} \pm 0.017$
S_4	$-0.266^{+0.063}_{-0.072} \pm 0.010$
S_5	$-0.323^{+0.063}_{-0.072} \pm 0.009$
A_FB	$0.305^{+0.049}_{-0.048} \pm 0.013$
S_7	$0.044^{+0.073}_{-0.072} \pm 0.013$
S_8	$0.013^{+0.071}_{-0.070} \pm 0.005$
S_9	$-0.094^{+0.065}_{-0.067} \pm 0.004$

$1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.690^{+0.035}_{-0.036} \pm 0.017$
S_3	$0.012^{+0.038}_{-0.038} \pm 0.004$
S_4	$-0.155^{+0.057}_{-0.056} \pm 0.004$
S_5	$-0.023^{+0.050}_{-0.049} \pm 0.005$
A_FB	$-0.075^{+0.032}_{-0.034} \pm 0.007$
S_7	$-0.077^{+0.050}_{-0.049} \pm 0.006$
S_8	$0.028^{+0.058}_{-0.057} \pm 0.008$
S_9	$-0.064^{+0.042}_{-0.041} \pm 0.004$

$15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
F_L	$0.344^{+0.028}_{-0.030} \pm 0.008$
S_3	$-0.163^{+0.033}_{-0.033} \pm 0.009$
S_4	$-0.284^{+0.038}_{-0.041} \pm 0.007$
S_5	$-0.325^{+0.036}_{-0.037} \pm 0.009$
A_FB	$0.355^{+0.027}_{-0.027} \pm 0.009$
S_7	$0.048^{+0.043}_{-0.043} \pm 0.006$
S_8	$0.028^{+0.044}_{-0.045} \pm 0.003$
S_9	$-0.053^{+0.039}_{-0.039} \pm 0.002$

Table 5.19: Results for the *CP*-averaged observables S_i.

$0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$0.006^{+0.064}_{-0.065} \pm 0.005$
A_4	$-0.068^{+0.071}_{-0.073} \pm 0.009$
A_5	$0.001^{+0.061}_{-0.059} \pm 0.018$
A_6	$0.122^{+0.076}_{-0.075} \pm 0.011$
A_7	$0.076^{+0.061}_{-0.060} \pm 0.006$
A_8	$-0.031^{+0.074}_{-0.074} \pm 0.007$
A_9	$0.030^{+0.062}_{-0.061} \pm 0.004$

$1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$0.042^{+0.097}_{-0.087} \pm 0.005$
A_4	$0.235^{+0.125}_{-0.109} \pm 0.005$
A_5	$-0.114^{+0.099}_{-0.105} \pm 0.009$
A_6	$0.037^{+0.102}_{-0.091} \pm 0.016$
A_7	$-0.087^{+0.091}_{-0.093} \pm 0.004$
A_8	$-0.044^{+0.108}_{-0.113} \pm 0.005$
A_9	$-0.004^{+0.092}_{-0.098} \pm 0.005$

$2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$-0.111^{+0.087}_{-0.109} \pm 0.006$
A_4	$-0.007^{+0.130}_{-0.135} \pm 0.007$
A_5	$-0.005^{+0.107}_{-0.106} \pm 0.008$
A_6	$0.022^{+0.115}_{-0.096} \pm 0.010$
A_7	$-0.032^{+0.109}_{-0.115} \pm 0.005$
A_8	$-0.071^{+0.124}_{-0.131} \pm 0.006$
A_9	$-0.228^{+0.114}_{-0.152} \pm 0.007$

$4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$-0.173^{+0.070}_{-0.079} \pm 0.006$
A_4	$-0.168^{+0.086}_{-0.085} \pm 0.008$
A_5	$-0.059^{+0.071}_{-0.073} \pm 0.011$
A_6	$-0.023^{+0.082}_{-0.075} \pm 0.005$
A_7	$0.041^{+0.083}_{-0.082} \pm 0.004$
A_8	$0.004^{+0.093}_{-0.095} \pm 0.005$
A_9	$0.062^{+0.078}_{-0.072} \pm 0.004$

$6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$0.064^{+0.067}_{-0.064} \pm 0.011$
A_4	$-0.037^{+0.073}_{-0.073} \pm 0.011$
A_5	$0.129^{+0.067}_{-0.066} \pm 0.012$
A_6	$0.047^{+0.062}_{-0.060} \pm 0.011$
A_7	$0.035^{+0.065}_{-0.067} \pm 0.003$
A_8	$-0.043^{+0.070}_{-0.069} \pm 0.006$
A_9	$0.110^{+0.061}_{-0.060} \pm 0.005$

$11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$0.132^{+0.075}_{-0.073} \pm 0.005$
A_4	$-0.100^{+0.082}_{-0.077} \pm 0.009$
A_5	$0.027^{+0.077}_{-0.076} \pm 0.010$
A_6	$0.024^{+0.069}_{-0.067} \pm 0.013$
A_7	$-0.008^{+0.073}_{-0.073} \pm 0.005$
A_8	$0.014^{+0.075}_{-0.073} \pm 0.005$
A_9	$-0.057^{+0.057}_{-0.059} \pm 0.006$

$15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$-0.034^{+0.056}_{-0.055} \pm 0.007$
A_4	$-0.071^{+0.064}_{-0.064} \pm 0.008$
A_5	$-0.076^{+0.065}_{-0.063} \pm 0.010$
A_6	$-0.085^{+0.062}_{-0.060} \pm 0.012$
A_7	$-0.105^{+0.058}_{-0.059} \pm 0.005$
A_8	$0.048^{+0.063}_{-0.063} \pm 0.003$
A_9	$0.091^{+0.059}_{-0.059} \pm 0.004$

$17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$-0.056^{+0.075}_{-0.073} \pm 0.017$
A_4	$-0.071^{+0.073}_{-0.073} \pm 0.011$
A_5	$0.008^{+0.073}_{-0.075} \pm 0.010$
A_6	$-0.127^{+0.080}_{-0.076} \pm 0.018$
A_7	$0.047^{+0.070}_{-0.069} \pm 0.013$
A_8	$0.022^{+0.072}_{-0.073} \pm 0.005$
A_9	$0.043^{+0.066}_{-0.067} \pm 0.005$

$1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$-0.072^{+0.038}_{-0.038} \pm 0.004$
A_4	$0.012^{+0.057}_{-0.056} \pm 0.005$
A_5	$-0.044^{+0.049}_{-0.047} \pm 0.005$
A_6	$0.020^{+0.061}_{-0.060} \pm 0.009$
A_7	$-0.045^{+0.050}_{-0.050} \pm 0.006$
A_8	$-0.047^{+0.058}_{-0.057} \pm 0.008$
A_9	$-0.033^{+0.040}_{-0.042} \pm 0.004$

$15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
A_3	$-0.035^{+0.043}_{-0.042} \pm 0.010$
A_4	$-0.079^{+0.047}_{-0.048} \pm 0.008$
A_5	$-0.035^{+0.047}_{-0.047} \pm 0.010$
A_6	$-0.110^{+0.052}_{-0.051} \pm 0.013$
A_7	$-0.040^{+0.045}_{-0.044} \pm 0.006$
A_8	$0.025^{+0.048}_{-0.047} \pm 0.003$
A_9	$0.061^{+0.043}_{-0.044} \pm 0.002$

Table 5.20: Results for the *CP* asymmetries A_i.

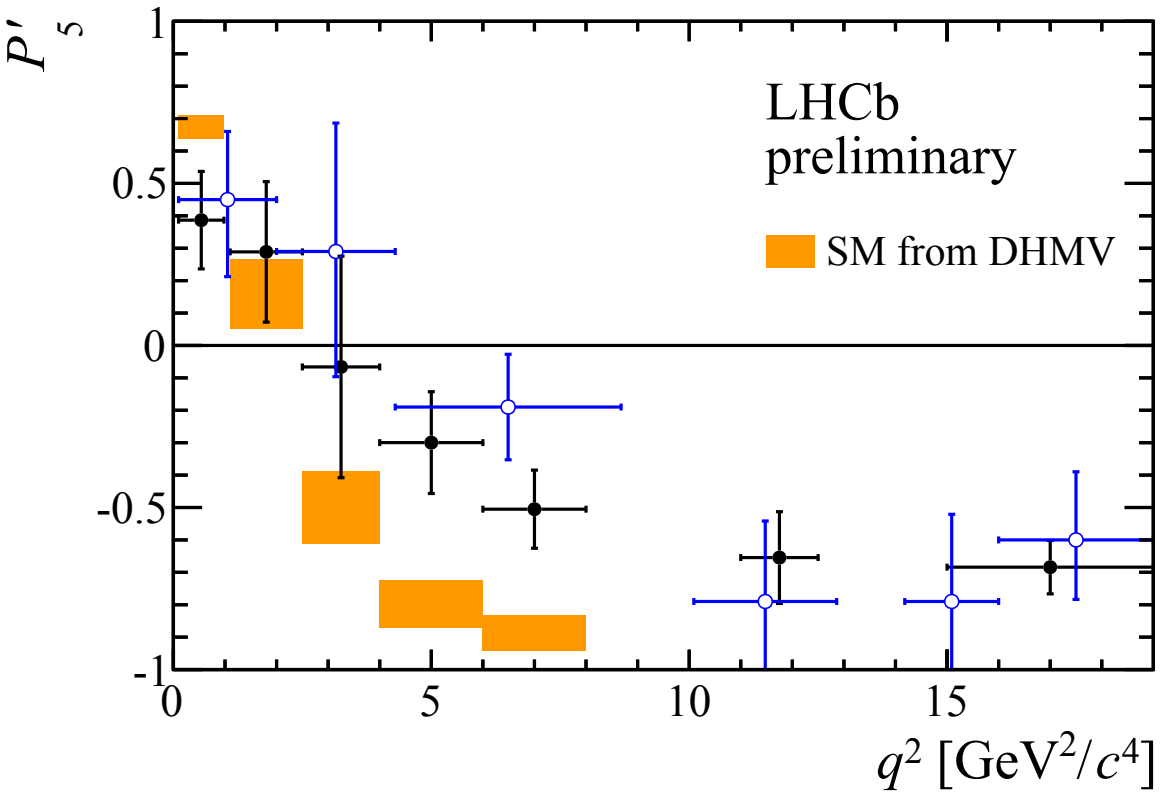


Figure 5.13: The P'_5 observable in bins of q^2. The yellow boxes are the SM prediction from Ref. [17]. In blue dots, the results from the previous analysis [10] are superposed for comparison.

$0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$-0.099^{+0.168}_{-0.163} \pm 0.014$
P_2	$-0.003^{+0.051}_{-0.052} \pm 0.007$
P_3	$0.113^{+0.079}_{-0.079} \pm 0.006$
P_4'	$0.185^{+0.158}_{-0.154} \pm 0.023$
P_5'	$0.387^{+0.132}_{-0.133} \pm 0.052$
P_6'	$0.034^{+0.134}_{-0.135} \pm 0.015$
P_8'	$0.180^{+0.174}_{-0.169} \pm 0.007$

$1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$-0.451^{+0.519}_{-0.636} \pm 0.038$
P_2	$-0.373^{+0.146}_{-0.199} \pm 0.027$
P_3	$0.350^{+0.330}_{-0.254} \pm 0.015$
P_4'	$-0.163^{+0.232}_{-0.240} \pm 0.021$
P_5'	$0.289^{+0.220}_{-0.202} \pm 0.023$
P_6'	$-0.463^{+0.202}_{-0.221} \pm 0.012$
P_8'	$-0.208^{+0.224}_{-0.270} \pm 0.024$

$2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$0.571^{+2.404}_{-1.714} \pm 0.045$
P_2	$-0.636^{+0.444}_{-1.735} \pm 0.015$
P_3	$0.745^{+2.587}_{-0.861} \pm 0.030$
P_4'	$-0.713^{+0.410}_{-1.305} \pm 0.024$
P_5'	$-0.066^{+0.343}_{-0.364} \pm 0.023$
P_6'	$0.205^{+0.962}_{-0.341} \pm 0.013$
P_8'	$0.091^{+0.650}_{-0.432} \pm 0.025$

$4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$0.180^{+0.364}_{-0.348} \pm 0.027$
P_2	$0.042^{+0.088}_{-0.087} \pm 0.011$
P_3	$0.083^{+0.187}_{-0.184} \pm 0.023$
P_4'	$-0.448^{+0.169}_{-0.172} \pm 0.020$
P_5'	$-0.300^{+0.158}_{-0.159} \pm 0.023$
P_6'	$-0.032^{+0.167}_{-0.166} \pm 0.007$
P_8'	$0.342^{+0.188}_{-0.185} \pm 0.009$

$6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$-0.199^{+0.281}_{-0.275} \pm 0.025$
P_2	$0.241^{+0.061}_{-0.062} \pm 0.013$
P_3	$0.057^{+0.148}_{-0.139} \pm 0.013$
P_4'	$-0.599^{+0.131}_{-0.135} \pm 0.010$
P_5'	$-0.505^{+0.122}_{-0.122} \pm 0.024$
P_6'	$-0.095^{+0.135}_{-0.135} \pm 0.011$
P_8'	$-0.171^{+0.142}_{-0.143} \pm 0.006$

$11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$-0.745^{+0.207}_{-0.230} \pm 0.015$
P_2	$0.418^{+0.053}_{-0.046} \pm 0.005$
P_3	$0.007^{+0.141}_{-0.138} \pm 0.010$
P_4'	$-0.567^{+0.169}_{-0.187} \pm 0.014$
P_5'	$-0.655^{+0.147}_{-0.160} \pm 0.015$
P_6'	$-0.282^{+0.146}_{-0.151} \pm 0.007$
P_8'	$-0.015^{+0.145}_{-0.142} \pm 0.005$

$15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$-0.436^{+0.134}_{-0.147} \pm 0.018$
P_2	$0.421^{+0.042}_{-0.035} \pm 0.005$
P_3	$0.029^{+0.082}_{-0.084} \pm 0.006$
P_4'	$-0.672^{+0.113}_{-0.151} \pm 0.016$
P_5'	$-0.662^{+0.109}_{-0.127} \pm 0.017$
P_6'	$0.127^{+0.119}_{-0.122} \pm 0.006$
P_8'	$0.007^{+0.125}_{-0.129} \pm 0.005$

$17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$-0.581^{+0.225}_{-0.263} \pm 0.037$
P_2	$0.314^{+0.046}_{-0.048} \pm 0.007$
P_3	$0.145^{+0.107}_{-0.102} \pm 0.008$
P_4'	$-0.556^{+0.133}_{-0.156} \pm 0.016$
P_5'	$-0.676^{+0.133}_{-0.152} \pm 0.017$
P_6'	$0.092^{+0.148}_{-0.152} \pm 0.025$
P_8'	$0.027^{+0.147}_{-0.147} \pm 0.009$

$1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$0.080^{+0.248}_{-0.245} \pm 0.044$
P_2	$-0.162^{+0.072}_{-0.073} \pm 0.010$
P_3	$0.205^{+0.135}_{-0.134} \pm 0.017$
P_4'	$-0.336^{+0.124}_{-0.122} \pm 0.012$
P_5'	$-0.049^{+0.107}_{-0.108} \pm 0.014$
P_6'	$-0.166^{+0.108}_{-0.108} \pm 0.021$
P_8'	$0.060^{+0.122}_{-0.124} \pm 0.009$

$15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$	
Obs.	Value
P_1	$-0.497^{+0.102}_{-0.099} \pm 0.027$
P_2	$0.361^{+0.025}_{-0.026} \pm 0.010$
P_3	$0.081^{+0.060}_{-0.059} \pm 0.005$
P_4'	$-0.597^{+0.080}_{-0.085} \pm 0.015$
P_5'	$-0.684^{+0.078}_{-0.081} \pm 0.020$
P_6'	$0.101^{+0.090}_{-0.092} \pm 0.011$
P_8'	$0.059^{+0.094}_{-0.093} \pm 0.008$

Table 5.21: Results for the *CP*-averaged observables $P_i^{(\prime)}$.

5.3 Theoretical interpretations of the results.

With the observation of a local discrepancy in the P_5' observable, many theoretical interpretations have been proposed (see for example [18], [19], [20], [21]). This discrepancy is actually not the only one within the electroweak penguin decays. Some other analysis have shown anomalies. For example the branching ratios observed in the $b \to s$ transitions: $B^+ \to K^{(*)}\mu^+\mu^-$, $B_s^0 \to \phi\mu^+\mu^-$ and $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ (see Ref. [22], [23], [24]) are systematically lower than the expectations by about one to three standard deviations, depending on the analysis. Also, the ratio R_K measured by LHCb [25] is 2.6 standard deviations away from the Standard Model prediction. Each of these results has the tendency to indicate the presence of a negative C_9^{NP}, affecting the muon channel in the case of R_K, though none of them by itself provide a clear evidence for new physics.

In order to interpret altogether these and other results concerning $b \to s$ transitions, a global fit providing a model-independent analysis has been performed. In the framework of the operator product expansion, described in Section 1.2.1, this global fit allows to set limits for the Wilson coefficients C_7, C_9 and C_{10} [21]. A χ^2 function, providing the compatibility between the model and the data, is minimised in different configurations, where different sets of Wilson coefficients are allowed to have a new physics contribution. The resulting χ^2 values are compared among them and with the one obtained assuming only the Standard Model. The new physics dependences are encoded in the Wilson coefficients, $C_i^{NP} = C_i - C_i^{\mathrm{SM}}$. This global fit makes use of 88 measurements of 76 different observables from the LHCb, ATLAS, CMS, BaBar, Belle and CDF collaborations. In particular it includes:

- the angular observables of $B^0 \to K^{*0}\mu^+\mu^-$;
- the differential branching ratios as function of q^2 of many decays : $B^0 \to K^{*0}\mu^+\mu^-$, $B^{*-} \to K^{*-}\mu^+\mu^-$, $B^- \to K^-\mu^+\mu^-$, $B_s \to \phi\mu^+\mu^-$, $B_s \to \mu^+\mu^-$, $B^0 \to K^{*0}\gamma$, $B^- \to K^{*-}\gamma$, $B \to X_s\gamma$, $B \to X_s\mu^+\mu^-$

The best fit is obtained when assuming new physics in C_9 only, with a $C_9^{\mathrm{NP}} = -1.07$, corresponding to a 3.7σ deviation from the Standard Model, or when $C_9^{\mathrm{NP}} = -C_{10}^{\mathrm{NP}} = -0.53$, with a pull deviating of 3.1σ from the SM. The results of this fit are shown in Figure. 6.16

To explain the anomaly in P_5' and in the global fit results, two main different interpretations have been advanced: either the presence of a new physics interaction is advocated, or an underestimated hadronic effect, such as the contribution of charm loops, is claimed. They will be briefly described in the following.

5.3.1 New physics interpretations.

There are two kind of new physics models that could agree with the results of the best fit, contributing to the C_9 Wilson coefficient. The first type concerns model where a new interaction is mediated by a new neutral gauge boson, Z', at tree-level. The second type of models contemplate the existence of scalar or vector leptoquarks.

In the case of models with a new Z' boson, different new physics scenario are actually possible.

- The Z' could couple to leptons through the $L_\mu - L_\tau$ current, which has an impact only on C_9 but also introduces a violation of the lepton flavour universality (see [26], [27], [28]).

- The Z' boson is part of a composite Higgs model, with a partially composite muon, again introducing a violation of the lepton flavour universality (see [29]). This interpretation prefers the solution $C_9^{\rm NP} = -C_{10}^{\rm NP}$.
- The Z' appears in an extension of the Standard Model gauge group with an additional $U(1)'$ symmetry, involving both quarks and leptons, and also including a degree of lepton flavour universality violation (see [30]).

In the case of models with scalar or vector leptoquarks (see [31], [32], [33]), the leptoquark could have spin 0 or 1 and have various possible representations. For example, a scenario with a single leptoquark leads to $C_9^{\rm NP} = \pm C_{10}^{\rm NP}$, while multiple leptoquarks would be needed for including new physics effect in multiple Wilson coefficients.

5.3.2 Hadronic effects

Another explanation for the observed anomalies could be that the hadronic contributions of the charm-resonances have been underestimated (see [34]). Indeed, these charmonium effects are photon mediated, implying a vector like coupling to leptons, which corresponds to the C_9 Wilson coefficient. One can include inside the effective C_9 Wilson a coefficient for the charm loop as follows:

$$C_9^{(\prime)\rm eff} = C_9^{(\prime)\rm SM} + C_9^{(\prime)\rm NP} + \eta^{(\prime)} h(q^2)$$

where $h(q^2)$ is the charm vacuum polarisation. The functional form of $h(q^2)$ is taken from the low-recoil $c\bar{c}$ resonances and then extrapolated at low q^2. In this way, from the best fits of the angular observables measured by LHCb, one could just get as solution $\eta = \eta' = -1.25$ and $C_9^{(\prime)\rm NP} = 0$, without advocating the presence of new physics.

5.4 Future experimental tests

The LHC has just restarted to collect data, and will continue until the next shutdown period in 2018. Then, after un upgrade of the detector, the data taking will resume. It is clear that, with more data available, all the measurements related to rare decays will be improved, as most of their errors are of statistical rather than systematical nature. Other additional measurements of rare decays, not yet possible with the current dataset, will become available too. The whole set of these new or improved measurements will enter the global fit of $b \to s$ transitions. This will certainly help to clarify the global picture and hopefully push today dicrepancies, if genuine, to evidences for new physics.

There are certain measurements that have been proposed, some of which already possible on a relatively short term timescale, that could help to shed light on the puzzle.

For example, since the charm loop effect is lepton flavour universal, a confirmation of the lepton flavour non-universality via the measurement of the R_{K^*} ratio could disfavour this interpretation. I take part in this measurement, currently ongoing in LHCb, and I will describe it in the following chapter.

Another way to test the hadronic effect is to measured more precisely $C_9^{\rm NP}$ in different q^2 bins. A new physics effect should be q^2-independent, while the charm-loops should have a q^2 dependence. This will require more data, as those that will be collected during the second run of data taking.

Some measurements could also help to distinguish between the different NP hypothesis. For example, a more precise measurement of $\mathcal{B}(B_s \to \mu^+\mu^-)$ will constrain C_{10}, allowing to

distinguish between models with new physics affecting only $C_9^{\rm NP}$ and models with $C_9^{\rm NP} = -C_{10}^{\rm NP}$. Also, the precise measurements of the ratio of $B \to K^* \ell^+ \ell^-$ angular observables and of the R_{K^*} ratio could disentangle between the Z' boson models or the leptoquark models. Indeed, for the NP model in Ref. [30], an equality in the ratio, $R_K = R_{K^*} = R_{X_S} = ...$ is expected, while $R_K = R_\eta$ and $R_{K^*} = R_\phi$ are preferred by Ref. [33].

REFERENCES

[1] $B^0 \to K^{*0}\mu^+\mu^-$ WG, $B^0 \to K^{*0}\mu^+\mu^-$ *selection*, LHCb-INT-2013-058, 2013.

[2] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.

[3] LHCb collaboration, R. Aaij *et al.*, *Measurement of the* $\bar{B}^0_{(s)} - B^0_{(s)}$ *production asymmetry in 7 TeV pp collisions*,arXiv:1408.0275, submitted to Phys. Lett. B.

[4] S. Stahl *et al.*, *Search for CP violation in* $D^0 \to K^-K^+, \pi^-\pi^+$ *using semileptonic B decays on 3* fb^-1,.

[5] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in* $D^0 \to K^-K^+$ *and* $D^0 \to \pi^-\pi^+$ *decays*, JHEP **07** (2014) 041, arXiv:1405.2797.

[6] BaBar, B. Aubert *et al.*, *Angular Distributions in the Decays B —> K* l+ l-*, Phys. Rev. **D79** (2009) 031102, arXiv:0804.4412.

[7] Belle, J. T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for B —> K(*)l+l-*, Phys. Rev. Lett. **103** (2009) 171801, arXiv:0904.0770.

[8] CDF, T. Aaltonen *et al.*, *Measurements of the Angular Distributions in the Decays* $B \to K^{(*)}\mu^+\mu^-$ *at CDF*, Phys. Rev. Lett. **108** (2012) 081807, arXiv:1108.0695.

[9] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay* $B^0 \to K^{*0}\mu^+\mu^-$, JHEP **08** (2013) 131, arXiv:1304.6325.

[10] LHCb, R. Aaij *et al.*, *Measurement of Form-Factor-Independent Observables in the Decay* $B^0 \to K^{*0}\mu^+\mu^-$, Phys. Rev. Lett. **111** (2013) 191801, arXiv:1308.1707.

[11] ATLAS Collaboration, *Angular Analysis of* $B_d \to K^{*0}\mu^+\mu^-$ *with the ATLAS Experiment*, Tech. Rep. ATLAS-CONF-2013-038, CERN, Geneva, Apr, 2013.

[12] CMS, S. Chatrchyan *et al.*, *Angular analysis and branching fraction measurement of the decay* $B^0 \to K^{*0}\mu^+\mu^-$, Phys. Lett. **B727** (2013) 77, arXiv:1308.3409.

[13] C. Bobeth, G. Hiller, and D. van Dyk, *More Benefits of Semileptonic Rare B Decays at Low Recoil: CP Violation*, JHEP **07** (2011) 067, arXiv:1105.0376.

[14] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, *Optimizing the basis of B->K*ll observables in the full kinematic range*, JHEP **05** (2013) 137, arXiv:1303.5794.

[15] CMS Collaboration, V. Khachatryan, A. M. Sirunyan *et al.*, *Angular analysis of the decay* $B^0 \to K^{*0}\mu^+\mu^-$ *from pp collisions at* $\sqrt{s} = 8$ *TeV*, Tech. Rep. arXiv:1507.08126. CMS-BPH-13-010. CERN-PH-EP-2015-178, CERN, Geneva, Jul, 2015. Comments: Submitted to Phys. Lett. B.

[16] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \to V\ell^+\ell^-$ in the Standard Model from Light-Cone Sum Rules*, (2015) arXiv:1503.0553.

[17] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *On the impact of power corrections in the prediction of $B \to K^*\mu^+\mu^-$ observables*, JHEP **12** (2014) 125, arXiv:1407.8526.

[18] S. Descotes-Genon, J. Matias, and J. Virto, *Understanding the $B \to K^*\mu^+\mu^-$ Anomaly*, Phys. Rev. **D88** (2013) 074002, arXiv:1307.5683.

[19] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *QCD uncertainties in the prediction of $B \to K^*mu^+mu^-$ observables*, (2014) arXiv:1411.0922.

[20] T. Hurth, F. Mahmoudi, and S. Neshatpour, *Global fits to $b \to s\ell\ell$ data and signs for lepton non-universality*, JHEP **12** (2014) 053, arXiv:1410.4545.

[21] W. Altmannshofer and D. M. Straub, *Implications of $b \to s$ measurements*, in *50th Rencontres de Moriond on EW Interactions and Unified Theories La Thuile, Italy, March 14-21, 2015*, 2015. arXiv:1503.6199.

[22] LHCb, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays*, JHEP **06** (2014) 133, arXiv:1403.8044.

[23] LHCb, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$*, (2015) arXiv:1506.0877.

[24] LHCb, R. Aaij *et al.*, *Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays*, JHEP **06** (2015) 115, arXiv:1503.0713.

[25] (LHCb Collaboration), R. Aaij *et al.*, *Test of Lepton Universality Using $B^+ \to K^+\ell^+\ell^-$ Decays*, Phys. Rev. Lett. **113** (2014) 151601.

[26] A. Crivellin, G. D'Ambrosio, and J. Heeck, *Explaining $h \to \mu^\pm\tau^\mp$, $B \to K^*\mu^+\mu^-$ and $B \to K\mu^+\mu^-/B \to Ke^+e^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$*, Phys. Rev. Lett. **114** (2015) 151801, arXiv:1501.0099.

[27] A. Crivellin, G. D'Ambrosio, and J. Heeck, *Addressing the LHC flavor anomalies with horizontal gauge symmetries*, Phys. Rev. **D91** (2015), no. 7 075006, arXiv:1503.0347.

[28] W. Altmannshofer and D. M. Straub, *New physics in $b \to s$ transitions after LHC run 1*, (2014) arXiv:1411.3161.

[29] C. Niehoff, P. Stangl, and D. M. Straub, *Violation of lepton flavour universality in composite Higgs models*, Phys. Lett. **B747** (2015) 182, arXiv:1503.0386.

[30] A. Celis, J. Fuentes-Martin, M. Jung, and H. Serodio, *Family nonuniversal Z? models with protected flavor-changing interactions*, Phys. Rev. **D92** (2015), no. 1 015007, arXiv:1505.0307.

[31] S. Sahoo and R. Mohanta, *Scalar leptoquarks and the rare B meson decays*, Phys. Rev. **D91** (2015), no. 9 094019, arXiv:1501.5193.

[32] S. Biswas, D. Chowdhury, S. Han, and S. J. Lee, *Explaining the lepton non-universality at the LHCb and CMS within a unified framework*, JHEP **02** (2015) 142, arXiv:1409.0882.

[33] G. Hiller and M. Schmaltz, *Diagnosing lepton-nonuniversality in $b \to s\ell\ell$*, JHEP **02** (2015) 055, arXiv:1411.4773.

[34] J. Lyon and R. Zwicky, *Resonances gone topsy turvy - the charm of QCD or new physics in $b \to s\ell^+\ell^-$?*, (2014) arXiv:1406.0566.

Chapter 6

R_{K^*} analysis at low q^2

Sommaire

The first LHCb measurement of the R_{K^*} ratio is well advanced and currently being finalised. The measurement is performed in three regions of q^2 and the analysis strategy is presented in Section 6.1. The rest of the chapter focuses on the measurement in the low q^2 region. As this measurement relies on the use of Monte Carlo simulations, it is crucial to apply corrections in order to correctly reproduces the data behaviour, as described in Section 6.2. These corrected samples are indeed used for a dedicated events selection, detailed in Section 6.3, as well as for the evaluation of the different components of the efficiency, listed in Section 6.4. Another key ingredient of the measurement is the determination of the signal yields, performed by the invariant mass fit procedure presented in Section 6.5. The systematics studies are shown in Section 6.6. All the analysis procedure is validated by the measurement of the $R_{J/\psi}$ ratio in Section. 6.7. At the time of writing, the result of the measurement is not known yet, as the analysis is still under review by other members of the LHCb collaboration. The whole analysis, including the measurement of the R_{K^*} ratio in the central and high-q^2 region, is described extensively in Ref. [1].

6.1 Analysis strategy

The R_{K^*} ratio is measured by reconstructing the $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ rare channels with the K^{*0} decaying into a kaon and a pion of opposite charges. In order to reduce the systematic uncertainties, the measurement is performed as a double ratio with respect to the decays reaching the same final states via a J/ψ resonance, $B^0 \to K^{*0}J/\psi\,(\to \ell^+\ell^-)$, also referred to as "charmonium" or "resonant" modes:

$$R_{K^*} = \frac{N_{B^0\to K^{*0}\mu^+\mu^-}}{N_{B^0\to K^{*0}J/\psi\,(\to\mu^+\mu^-)}} \cdot \frac{N_{B^0\to K^{*0}J/\psi\,(\to e^+e^-)}}{N_{B^0\to K^{*0}e^+e^-}} \cdot \frac{\varepsilon_{B^0\to K^{*0}J/\psi\,(\to\mu^+\mu^-)}}{\varepsilon_{B^0\to K^{*0}\mu^+\mu^-}} \cdot \frac{\varepsilon_{B^0\to K^{*0}e^+e^-}}{\varepsilon_{B^0\to K^{*0}J/\psi\,(\to e^+e^-)}}. \tag{6.1}$$

Indeed, using the relative efficiencies between the rare and resonant modes allows to cancel many systematic effects, as they wuld affect both the numerator and the denominator, resulting in a better control of the systematic uncertainties. In addition, the resonant modes are used as control samples, and they also provide a powerful test of the analysis procedure: as the new physics is expected not to affect the decays with the charmonium resonances, the ratio $R_{J/\psi}$ of the two $B^0 \to K^{*0}J/\psi\,(\to \ell^+\ell^-)$ channels is expected to be unity.

For brevity, in the following the rare channels are also referred to as "$\ell\ell$", or specifically "$\mu\mu$" and "ee", and the charmonium modes as "$J/\psi\,(\ell\ell)$", or "$J/\psi\,(\mu\mu)$" and "$J/\psi\,(ee)$". The invariant mass of the dilepton pair is used to distinguish among the rare and the resonant decays.

The full analysis is performed in three regions of q^2:

- **low-q^2**: below $1.1\,\mathrm{GeV}^2/c^4$, where the $b \to s\ell^+\ell^-$ process is dominated by the photon pole;
- **central-q^2**: between 1.1 and $6.0\,\mathrm{GeV}^2/c^4$;
- **high-q^2**: above $15\,\mathrm{GeV}^2/c^4$.

In order to fully include in the lower region the $B^0 \to K^{*0}\phi$ decay with $\phi \to \ell^+\ell^-$, which could dilute the effects of new physics, the boundary between the low and the central-q^2 regions is set to $1.1\,\mathrm{GeV}^2/c^4$. The upper bound of the central bin is chosen to be sufficiently away from the $J/\psi\,(ee)$ radiative tail, where predictions cannot be cleanly extracted. The region between 6 and $15\,\mathrm{GeV}^2/c^4$ is dominated by the presence of the J/ψ and $\psi(2S)$ resonances, which are used as control channels as stated before. The lower bound of the high-q^2 interval, where the signal in

the electron channel is still unobserved, is chosen to be sufficiently far from the $\psi(2S)$ resonance to avoid its contamination.

The following chapter will mainly focus on the analysis realised on this low-q^2 region, though many procedures are common with the other regions. Indeed, the low-q^2 region is particularly interesting. It is mainly dominating by the Wilson coefficient C_7, which is well measured and known to be not affected by new physics effects at the level we are searching for. For this reason, if the R_{K^*} ratio measured in the low-q^2 region will not surprisingly reveal a deviation from the Standard Model, still it will validate the whole analysis procedure and provide confidence in the measurements performed in the other two regions.

6.2 Data-simulation corrections

Since the training of the multivariate classifiers (see Section. 6.3.5) and the determination of most of the efficiency components (see Section. 6.4) are based on simulation, it is crucial that the simulation reproduces the data behaviour reliably. In particular, it is important that the detector occupancy and the kinematics of the final state particles match. The reconstructed 4-body invariant mass of the ee channels is strongly correlated to the detector occupancy as this directly affects the bremsstrahlung recovery algorithm. The hit multiplicity in the SPD detector is used as a proxy for the detector occupancy. Instead, the description of the kinematics of the decay is probed by looking at the transverse momentum spectrum of the B^0. Discrepancies in this distribution cause also the spectra of the final particles to differ and affect the efficiency determination, as this depends on the momentum of the final state particles.

The correction factors are determined using $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ decays, for which the signal peak is already visible in data after the pre-selection (see Section. 6.3), and applied to both muon and electron simulated samples, for which no differences are expected in the SPD and $p_T(B^0)$ distributions. The ${}_s\mathcal{P}lot$ technique [2] is used to statistically subtract the background and obtain pure $B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)$ signal distributions. The invariant mass of the B^0 is used as the control variable, after having constrained J/ψ vertex and mass to the PDG value. The Figure. 6.1 shows the fit to the 4-body invariant mass $m_{K\pi\mu\mu}$ of the $J/\psi\,(\mu\mu)$ candidates right after the pre-selection.

The discrepancy in the SPD hit multiplicity is corrected first. Then the B^0 transverse momentum distributions in data and simulation, reweighted for the SPD multiplicities, are compared. These distributions are presented in Figure. 6.2 and their ratios, which are used to reweight the simulation, are shown in Fig. 6.3. The weights for the SPD multiplicity are determined separately for 2011 and 2012, since the average number of interactions per bunch crossing was different in the two years. To limit statistical fluctuations, an adaptive binning is used, chosen to provide approximately the same number of simulated events in each bin. Since the transverse momentum of the B^0 is strongly affected by the electron bremsstrahlung, the corresponding correction is applied using the true B^0 p_{T}.

6.3 Selection at low q^2

As it was the case for the angular analysis described in the first part of this thesis, the selection of the $B^0 \to K^{*0} l^+ l^-$signal candidates in the low q^2 is made up of several steps. The first are the trigger selection and the "Stripping", described in the following Sections 6.3.1 and 6.3.2. Then a dedicated particle identification criteria and mass cuts are applied against peaking backgrounds, as described in Section 6.3.4. Finally a multivariate technique to reduce the amount of

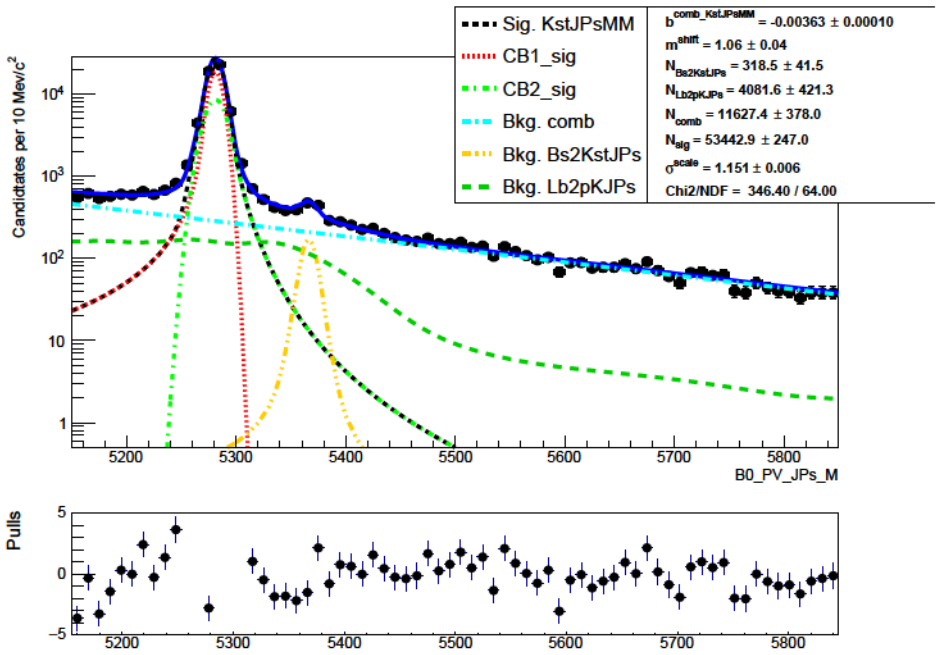


Figure 6.1: Fit to the $m_{K\pi\mu\mu}$ invariant mass of $B^0 \to K^{*0} J/\psi (\to \mu^+\mu^-)$ candidates after preselection.

combinatorial background is applied, described in detail in Section 6.3.5. The selection strategy for the $\mu\mu$ and ee channels are mainly based on the analysis in [3] and Ref. [4] respectively. Here we describe the electron channel selection that was not described before. For the muon channel, we put the accent on the need for a specific optimisation at low q^2 with respect to the selection developed for the angular analysis and on the resulting gain.

6.3.1 Trigger

The trigger requirements have been established in common for the three regions of q^2. For the $B^0 \to K^{*0}\mu^+\mu^-$ candidates, events are required to pass the same trigger requirements described in Sec 3.2, with the exception of the Hlt2SingleMuon line. This is a prescaled trigger, so its efficiency calculation is more complicated, and since it was adding a negligible number of signal events, it was discarded. Candidates of the $B^0 \to K^{*0}e^+e^-$ decay are required to pass the same trigger requirements used in the $B^0 \to K^{*0}e^+e^-$ angular analysis [4]. The trigger line required are listed in Tab. 6.1.

For the electron channel, previous analyses [4, 5] have noticed a dependence of the reconstructed B meson invariant mass from the L0 trigger line which had selected the event. Therefore, this analysis will be performed separately in three different trigger categories, defined as:

- **L0E**: events triggered by at least one of the electrons in the signal candidate (`L0Electron_TOS`);
- **L0H**: events triggered by at least one of the hadrons in the signal candidate and not by the electrons (`L0Hadron_TOS && !L0Electron_TOS`);
- **L0I**: events triggered by particles not in the signal candidate and not belonging to any of the two previous cases (`L0Global_TIS && !(L0Electron_TOS || L0Hadron_TOS)`).

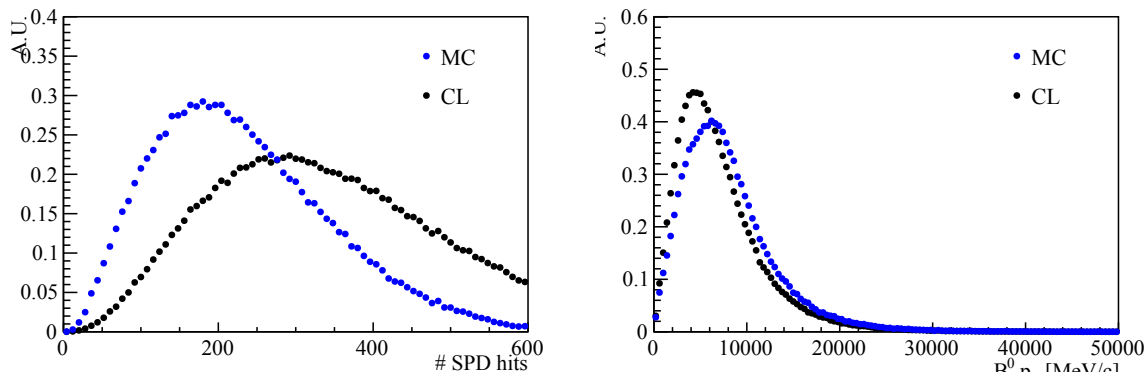


Figure 6.2: Distributions of the number of SPD hits (left) and the B^0 transverse momentum (right) in 2012 data (CL) and simulation (MC).

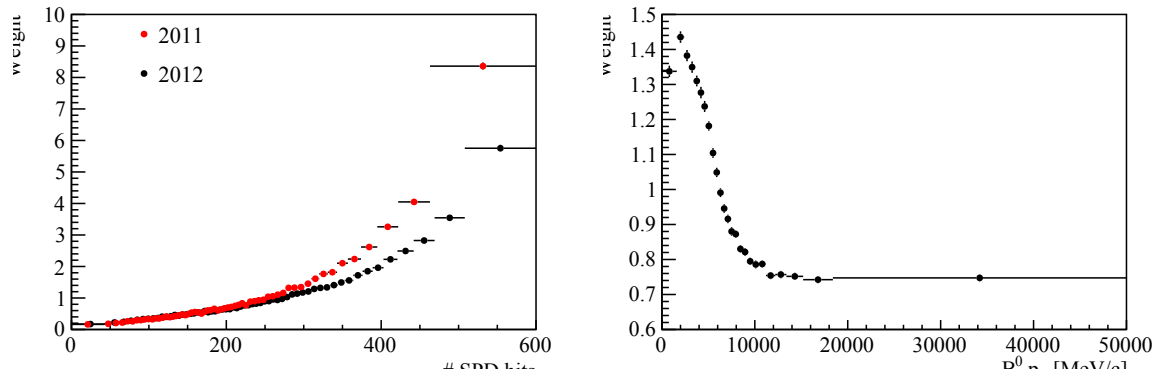


Figure 6.3: Ratios of simulated over real data distributions used to correct the Monte Carlo as a function of the number of SPD hits (left) and the B^0 transverse momentum (right).

Muon	**Electron**
L0Muon	L0Electron L0Hadron L0Global (TIS)
Hlt1TrackAllL0 Hlt1TrackMuon	Hlt1TrackAllL0
Hlt2Topo[2,4]BodyBBDT Hlt2TopoMu[2,4]BodyBBDT Hlt2DiMuonDetached	Hlt2Topo[2,4]BodyBBDT Hlt2TopoE[2,4]BodyBBDT

Table 6.1: Summary of the trigger lines used to select the $\mu\mu$ and the ee channels. All lines are required to be TOS, except when specified differently.

Most of the events are selected by the L0E category. The L0H is more efficient in the low-q^2 region, where the K^{*0} is more energetic than in the other regions. Since L0I is defined as independent of the signal candidate, the corresponding trigger efficiency cancels in the ratio between the rare and resonant modes.

6.3.2 Stripping

The $B^0 \to K^{*0}\mu^+\mu^-$ candidates are selected with the stripping line described in Section 3.3. Instead, the $B^0 \to K^{*0}e^+e^-$ candidates are selected with a BDT-based stripping line described in [4]. In this stripping line, first all cut outlined in Table 6.2 are applied. Then a BDT classifier is used with the variables listed in Table 6.3 as input. The BDT has been trained using a $B^0 \to K^{*0}e^+e^-$ Monte Carlo as signal proxy and upper mass sideband events, $m_{K\pi e^+e^-} > 5600$ MeV/c^2, as background proxy.

Particle	Condition
B^0	$\chi^2_{\text{Vertex}} < 16$, $\lvert m_{B^0} - 5280\rvert < 1000$ MeV/c^2, $\theta_{\text{flight}} > 0.999$
K^{*0}	$\chi^2_{\text{Vertex}} < 16$, $\lvert m_{K^{*0}} - 895.9\rvert < 150$ MeV/c^2
e^+e^- (J/ψ)	$\chi^2_{\text{Vertex}} < 16$, $m(e^+e^-) = [1,1200]$ ($[2200,4200]$) MeV/c^2
K	$p_T > 400$ MeV/c^2, $p > 3$ GeV/c, $\chi^2_{\text{Track}} < 3$, $\chi^2_{\text{IP}} > 4$, PIDK > -5, `GoshtProb <0.35`
π	$p_T > 250$ MeV/c^2, $p > 2$ GeV/c, $\chi^2_{\text{Track}} < 3$, $\chi^2_{\text{IP}} > 4$, PIDK < 10, `GoshtProb <0.35`
$e^\pm$	$p_T > 200$ MeV/c^2, $\chi^2_{\text{Track}} < 3$, $\chi^2_{\text{IP}} > 1$, PIDe > -2, `GoshtProb <0.5`

Table 6.2: List of cuts of the stripping line for the electron channel.

Particle	BDT Input Variables
B^0	p_T, χ^2_{IP}, χ^2_{FD}, θ_{flight}
K^{*0}	p_T, χ^2_{IP}, χ^2_{FD}
e^+e^- (J/ψ)	p_T, χ^2_{IP}, χ^2_{FD}
$K, \pi, e^\pm$	p_T, χ^2_{IP}

Table 6.3: Variables used to train the BDT for the electron channel stripping line.

6.3.3 Pre-selection

After the trigger and the stripping, the candidates have to pass additional criteria. Both $\mu\mu$ and ee channels tighten the invariant mass window for the K^{*0} to 895.9 ± 100.0 MeV/c^2. Specific requirements for the $\mu\mu$ channel are detailed explicitly in Table A.2 in Section 3.4, For the ee channel the $m(e^+e^-)$ invariant mass window to select $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ is [2200,3400] MeV/c^2 and [20,1100] MeV/c^2 for the $B^0 \to K^{*0}e^+e^-$ decay.

6.3.4 Peaking background

Several sources of exclusive backgrounds have been considered. Some of them are common to both $\mu\mu$ and ee channels and are described in Section 3.5.1-3.5.6. Additional peaking background for the ee channel are vetoed as described in the following.

6.3.4.1 The $B^0 \to D^- e^+ \nu$ veto cut

The decay $B^0 \to D^- e^+ \nu$, where the D^- decays semileptonically to $K^{*0}e^-\nu$, has a branching ratio four orders of magnitude larger than $B^0 \to K^{*0}e^+e^-$ and it might pass the selection cuts if the two neutrinos carries a very low momentum. To avoid this scenario, a cut is applied to the invariant mass of the $K^{*0}(\bar{K^{*0}})$ and the $e^-(e^+)$ above the D^- mass at 1900 MeV/c.

6.3.4.2 The $B^0 \to K^{*0}\gamma$ veto cut

The branching fraction of $B^0 \to K^{*0}\gamma$ has been measured to be $\mathcal{B} = (4.33 \pm 0.15) \times 10^{-5}$, and in the case where the photon converts into an electron and a positron, this decay will have similar characteristics as $B^0 \to K^{*0}e^+e^-$. To apply a veto for it, the reconstructed invariant mass window for the di-electron is chosen above 20 MeV/c, so to cut the photon pole, and the error on the reconstructed z-coordinate of the e^+e^- pair is required to satisfy the condition $\sigma_Z(e^+e^-) < 30$ mm, so to ensure that the e^+e^- pair are created at the B meson vertex.

6.3.4.3 The $B^0 \to K^{*0}\eta$ decay

The $B^0 \to K^{*0}\eta$ decay can be a source of background if the η decays into two photons which convert into a di-electron pair, or if the η makes a Dalitz pair. The first case is similar to the $B^0 \to K^{*0}\gamma$ decay and thus is vetoed by the cut previously described (see Section 6.3.4.2). For the second case, where the η makes a Dalitz pair, the contribution was estimated to be about 3.3 events, with a flat mass distribution between 4800 MeV/c and 5000 MeV/c, so this is considered negligible.

6.3.4.4 The $B^0 \to K^{*0}V(\to e^+e^-)$ decays with V $= \rho,\omega,\phi$

As for the muon channel, contributions of light resonances have been studied [6] and found to be negligible for ρ and ω. For the $\phi(1020)$ resonance, due to the Bremsstrahlung radiation of the electrons, a complete veto of the resonance will require a larger q^2 cut for the electron than for the $\mu\mu$ channel. Since only 2.6 ± 0.6 events are expected for the $B^0 \to K^{*0}\phi(\to e^+e^-)$ decay, we choose to fully include the $B^0 \to K^{*0}\phi$ for both the $\mu\mu$ and the ee channels in the low q^2 region, rather than put a veto on it. This justify the choice of upper limit for the ow q^2 bin.

6.3.5 Multivariate analysis

To reduce the combinatorial background a multivariate classifier is used. For the ee channel the same BDT developed for the angular analysis [4] has been used, while for the $\mu\mu$ channel a new classifier has been developed and optimised for the low q^2 region. Since in this case we do not need an unbiased angular distribution, the p_T distributions of the final states particles have been added as input variables. In addition, since the J/ψ control sample does not reproduce correctly the distributions in the low-q^2, a new proxy for the signal as been used for the training of the classifier. The description and the performance of this new classifier are presented in this section.

6.3.5.1 Input samples

As previously described in Section 3.6, the classifier need calibration samples as proxy for both signal and background events, in order to be trained.

The input signal sample was obtained from a $B^0 \to K^{*0}\mu^+\mu^-$ simulated sample in the [0.047-1.1] GeV/c^2 bin, which has been corrected for data-simulation differences has described in Section 6.2. The background sample was obtained from the 2011 and 2012 LHCb data. The events in the upper mass sideband ($m_{K\pi\mu\mu} > 5350\,\text{MeV}/c^2$) belonging to the low q^2 region ([0.047-1.1] GeV^2/c^4) have been selected as background proxy. In order to be able to use the full statistic recorded by LHCb in 2011 and 2012, the data sample has been split randomly into two sub-samples and two classifiers have been trained. Then, each of this classifiers has been applied on the data sub-sample not used for its training.

6.3.5.2 Classifier trainings

A BDT with the *gradientBoost* option was used as classifier. The input variables are similar to those in Section 3.6.3 and [4] with the addition of the p_T. They are listed in Table 6.4. The distribution for each input variable for the signal and background samples can be see in Figure 6.4. The BDT of the $B^0 \to K^{*0} e^+ e^-$ channes as in addition variables from the K^{*0} and $\ell\ell$, which have been removed in the muon channel due to their high correlations with the variables of the final state particles (see Figure E.2).

Particle	Variables
B^0	p_{T}, $\log(\chi^2_{\mathrm{IP}})$, $\log(\chi^2_{FD})$, $\log(\chi^2_{vtx})$, `DIRA`
K	p_{T}, isolation, $\log(\chi^2_{\mathrm{IP}})$
π	p_{T}, isolation, $\log(\chi^2_{\mathrm{IP}})$
$\mu^{\pm}$	p_{T}, isolation, $\log(\chi^2_{\mathrm{IP}})$

Table 6.4: Input variables used in the muon channel BDT classifier.

After the training, the response of the two BDTs are similar, as seen in Fig 6.5. The comparison with the classifier used for the $B^0 \to K^{*0}\mu^+\mu^-$ angular analysis, in Fig 6.6, shows an improvement in the signal efficiency for this BDT dedicated to the low q^2 region.

6.3.6 Optimisation of the BDT cut and particle identification

The BDT cut was optimised using a similar strategy for both the $\mu\mu$ and ee channel. The optimisation consist in maximising the signal significance for the rare modes, $S = \frac{N_S}{\sqrt{N_S+N_B}}$, where N_S (N_B) is the number of expected signal (background) events in the signal mass window ($\pm 50\,\mathrm{MeV}/c^2$ around the B^0 mass). The BDT cut optimisation is done simultaneously to the optimisation of the PID cuts for the final state particles. Thus an optimal combination of the BDT and PID cuts has been obtained for each channel and for the different trigger categories. The PID variables considered are:

- the kaon `ProbNNk`,
- the pion `ProbNNpi`,
- the leptons PID, PIDe or PIDmu depending on the decay channel

An example of this optimisation can be seen in Figure 6.7 and the complete set of optimised cuts are reported for both channel in Table 6.5

As the simulated event samples do not reproduce faithfully the PID variables distributions, the number of expected signal events N_S is obtained using the number of signal found in the data for the $B^0 \to J/\psi\, K^{*0}$ decay and scaled with the selection efficiencies and known branching ratios:

$$N_S(B^0 \to K^{*0}\ell^+\ell^-) = \frac{N_S(B^0 \to K^{*0} J/\psi(\to \ell^+\ell^-)) \times \epsilon^{MC}(B^0 \to K^{*0}\ell^+\ell^-) \times \mathcal{B}(B^0 \to K^{*0}\ell^+\ell^-)}{\epsilon^{MC}(B^0 \to J/\psi\, K^{*0}) \times \mathcal{B}(B^0 \to J/\psi\, K^{*0}) \times \mathcal{B}(J/\psi \to \ell^+\ell^-)}$$

The number of signal events in the $B^0 \to K^{*0} J/\psi(\to \ell^+\ell^-)$ decay is extracted using a fit of the B^0 invariant mass with the $\ell^+\ell^-$ pair invariant mass fixed to the J/ψ PDG central value. The selection efficiencies have been extracted from MC.

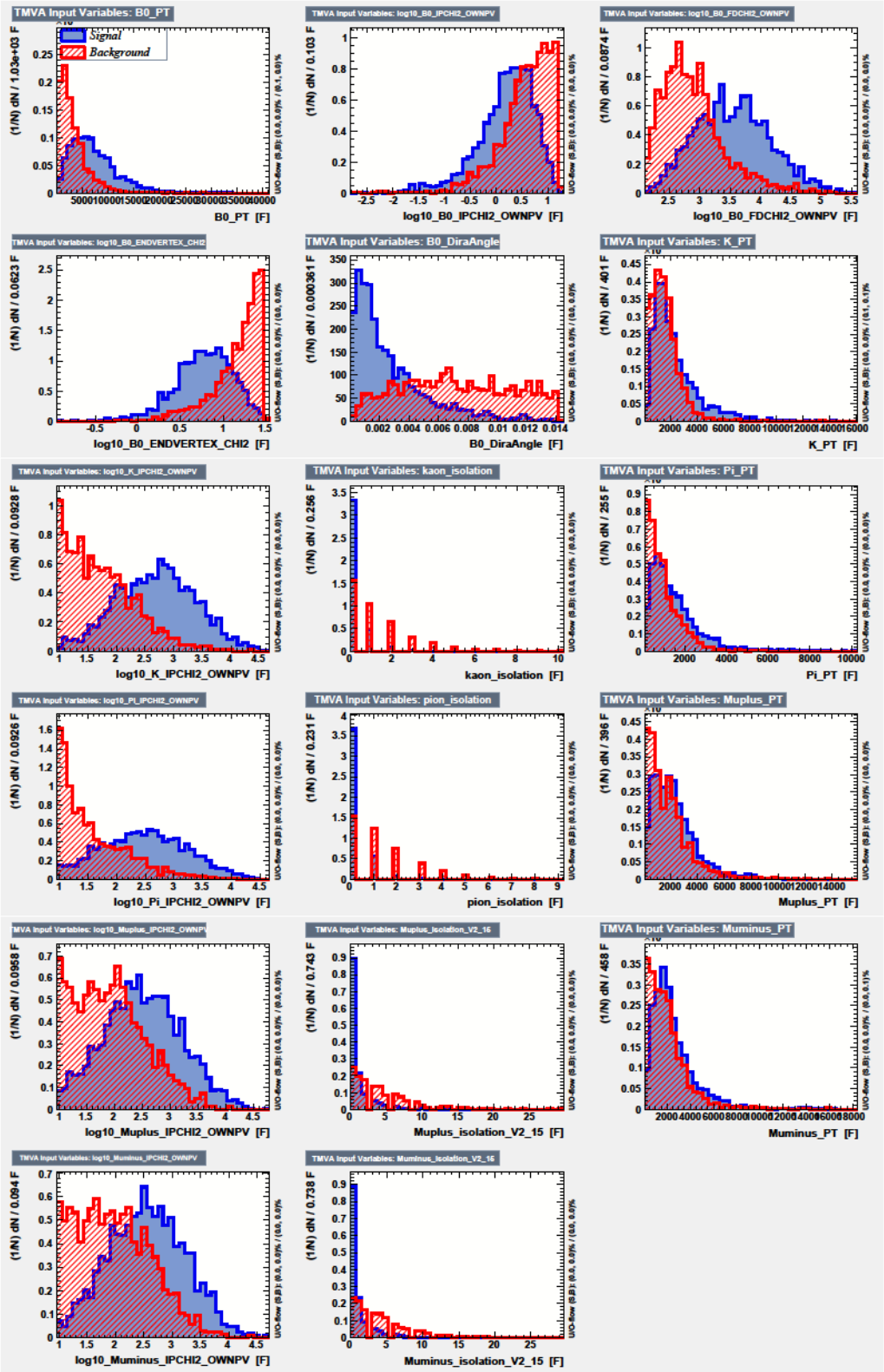


Figure 6.4: Distributions of the input variables used in the BDT classifier on the R_{K^*} analysis for the muon channel.

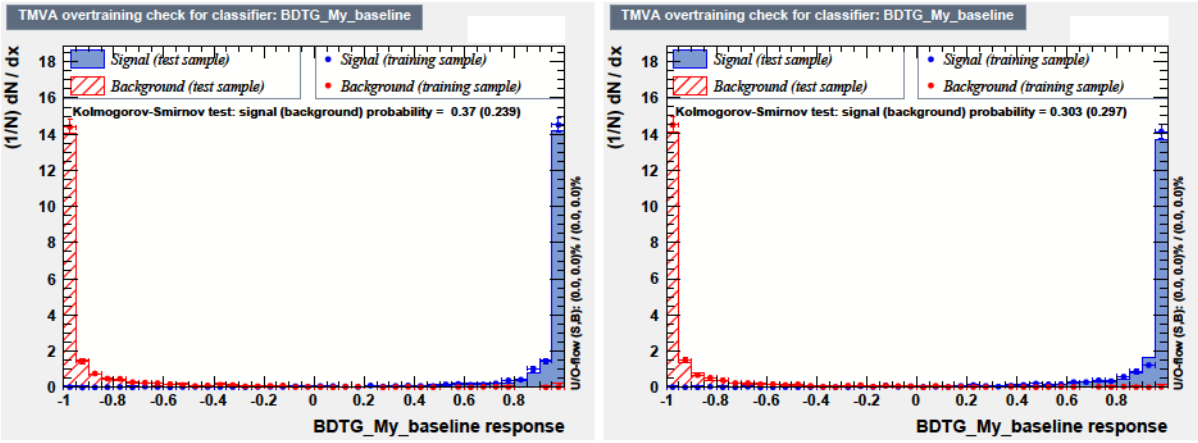


Figure 6.5: Responses of the BDTs for the two sub-samples used in the training, which are in good agreement. For each of them, the comparison of the distribution for the testing sample (shaded) and the training sample (dots) is also shown, demonstrating that there is no overtraining for both signal (blue) and background (red) samples.

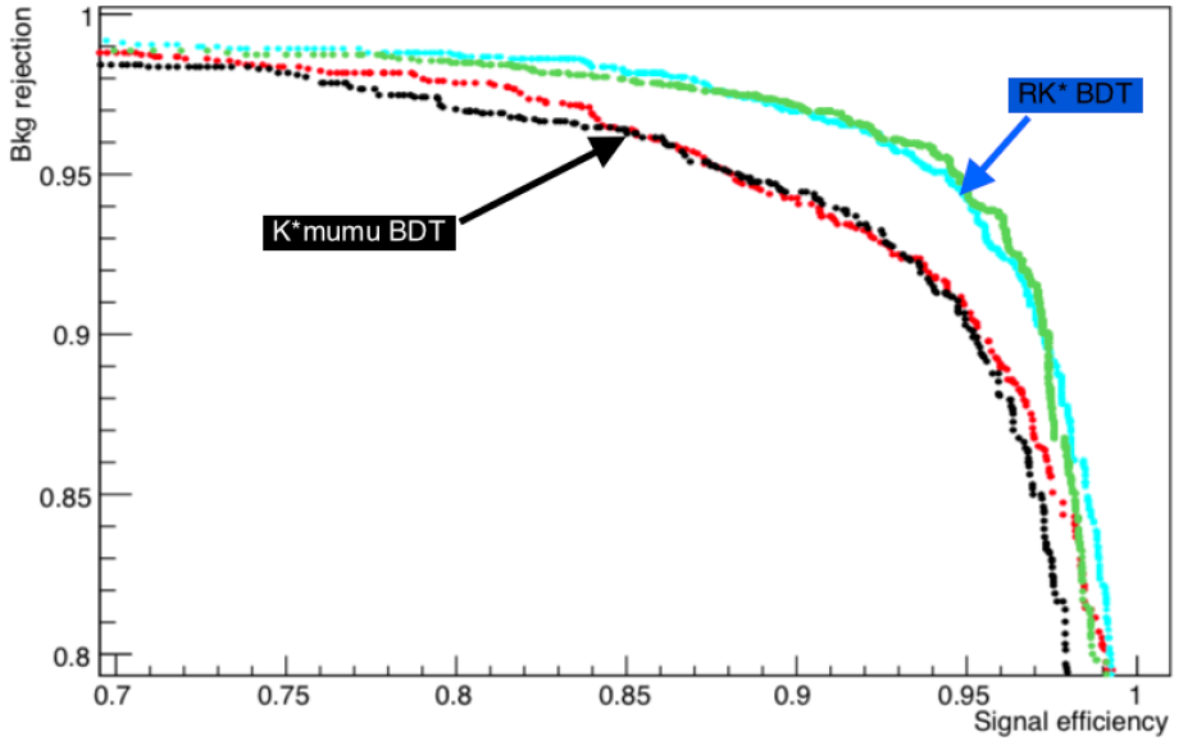


Figure 6.6: Signal efficiency versus background rejection for the two sub-samples (in green and cyan) in comparison with the BDT used for the $B^0 \to K^{*0}\mu^+\mu^-$ angular analysis applied to the same sub-samples (in black and red). The blue arrow show the working point of the new BDTs after the optimisation described in Section 6.3.6, to be compared with the working point of the BDT used for the $B^0 \to K^{*0}\mu^+\mu^-$ angular analysis (black arrow).

The background N_B expected in the blinded signal mass window has been extrapolated from a fit on the data mass-sidebands of the low q^2 region. These are: between 5100 and 5180 MeV/c^2

and between 5500 and 7000 MeV/c^2 for the muon channel; below 4800 and between 5500 and 6300 MeV/c^2 for the electron channel.

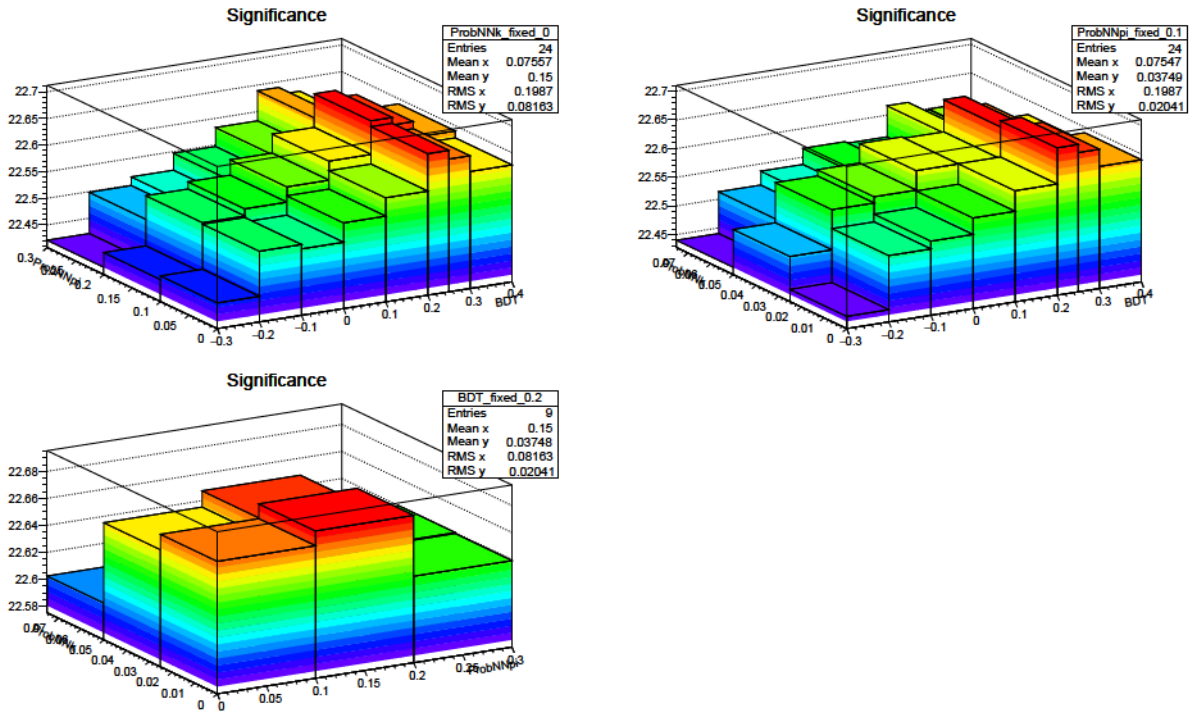


Figure 6.7: Example of results from the procedure maximizing a significance $\frac{N_S}{\sqrt{N_S+N_B}}$ in the μ channel for different combination of the cuts on the BDT, the K `ProbNNk` and the π `ProbNNpi` , with the $\mu^{\pm}$ PIDmu cut fixed to -3. Top left: BDT vs π ProbNNpi for K `ProbNNk` fixed to 0. Top right: BDT vs K ProbNNk for π `ProbNNpi` fixed to 0.1. Bottom left: K `ProbNNk` vs π ProbNNpi for BDT fixed to 0.2.

Channel		BDT	$\ell^{\pm}PID$	K `ProbNNk`	π `ProbNNpi`	$S/\sqrt{S+B}$
$B^0 \to K^{*0}e^+e^-$	L0Electron	0.88	1.2	0.05	0.2	7.5
	L0Hadron	0.94	1.2	0.05	0.2	3.9
	L0TIS	0.96	1.6	0.05	0.2	4.9
$B^0 \to K^{*0}\mu^+\mu^-$		0.2	0.0	0.0	0.1	22.7

Table 6.5: Optimized cuts and significances.

6.4 Evaluation of the analysis efficiencies

For each decay channel, the total efficiency of the analysis is calculated according to the formula:

$$\varepsilon^{tot} = \varepsilon^{\text{geom}} \cdot \varepsilon^{\text{reco|geom}} \cdot \varepsilon^{\text{trig|reco}} \cdot \varepsilon^{\text{PID|trig}} \cdot \varepsilon^{\text{MVA|PID}},$$

where $\varepsilon^{\text{geom}}$ corresponds to the efficiency of having all final state particles in the LHCb detector acceptance; $\varepsilon^{\text{reco|geom}}$ represents the reconstruction efficiency, given the geometrical efficiency; $\varepsilon^{\text{trig|reco}}$ corresponds to the trigger efficiency given the reconstruction efficiency; $\varepsilon^{\text{PID|trig}}$ repre-

sents the PID efficiency, given the trigger efficiency; $\varepsilon^{\text{MVA|PID}}$ corresponds to the efficiency of the MVA classifier, given the PID efficiency.

Reconstruction, trigger and MVA efficiencies are evaluated using the simulations, with the trigger efficiency for $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ being compared between data and simulation using the TISTOS method. The PID efficiency is calculated with the data-driven method described in Sec. 6.4.3.

The absolute efficiency for both rare and resonant channel are given in Tab. 6.6 and Tab. 6.7 respectively. The relative efficiencies of the rare over the resonant decay are given in Tab. 6.8. The rare to resonant efficiency double ratios $(\varepsilon_{ee}/\varepsilon_{J/\psi(ee)})/(\varepsilon_{\mu\mu}/\varepsilon_{J/\psi(\mu\mu)})$, given in Tab. 6.9, are the most relevant to the analysis as they enter directly the measurement of R_{K^*} (see equation 6.1).

ε	$\mu\mu$	ee		
		L0E	L0H	L0I
Geom	0.1727 ± 0.0005	0.1659 ± 0.0005		
Reco	0.0942 ± 0.0001	0.0504 ± 0.0001		
Trig	0.6076 ± 0.0004	0.1366 ± 0.0000	0.0518 ± 0.0000	0.1399 ± 0.0000
PID	0.9804 ± 0.0000	0.9106 ± 0.0000	0.8883 ± 0.0000	0.8892 ± 0.0000
MVA	0.9465 ± 0.0003	0.8759 ± 0.0007	0.7872 ± 0.0006	0.5360 ± 0.0017
Tot	0.0092 ± 0.0000	0.0009 ± 0.0000	0.0003 ± 0.0000	0.0006 ± 0.0000

Table 6.6: Absolute efficiencies for the rare ee and $\mu\mu$ channel in the low-q^2 region.

ε	$\mu\mu$	ee		
		L0E	L0H	L0I
Geom	0.1598 ± 0.00046	0.15934 ± 0.0004		
Reco	0.08875 ± 0.0001	0.0354 ± 0.0001		
Trig	0.7612 ± 0.0004	0.2560 ± 0.0004	0.0222 ± 0.0002	0.1159 ± 0.0004
PID	0.9810 ± 0.0000	0.9093 ± 0.0001	0.9065 ± 0.0004	0.8934 ± 0.0002
MVA	0.9047 ± 0.0003	0.8913 ± 0.0010	0.7700 ± 0.0046	0.5711 ± 0.0024
Tot	0.0094 ± 0.0000	0.0012 ± 0.0000	0.0001 ± 0.0000	0.0003 ± 0.0000

Table 6.7: Absolute efficiencies for the resonant channel, $J/\psi\,(\ell\ell)$ with the low-q^2 region.

ε	low-q^2			
	$\mu\mu$	ee		
		L0E	L0H	L0I
Geom	1.085 ± 0.018	1.041 ± 0.004		
Reco	1.061 ± 0.0011	1.424 ± 0.005		
Trig	0.798 ± 0.0039	0.5334 ± 0.0030	2.331 ± 0.030	1.207 ± 0.009
PID	0.999 ± 0.0000	1.0014 ± 0.0004	0.9799 ± 0.0010	0.9953 ± 0.0005
MVA	1.046 ± 0.0025	0.9827 ± 0.0037	1.0223 ± 0.0105	0.9386 ± 0.0094
Tot	0.960 ± 0.0108	0.7778 ± 0.0089	3.4603 ± 0.0655	1.6709 ± 0.0270

Table 6.8: Relative efficiencies of the rare over the resonant ($\varepsilon_{rel} = \varepsilon_{\ell\ell}/\varepsilon_{J/\psi}$) ee and $\mu\mu$ channels in the low q^2 region.

ε	Low-q^2		
	L0E	L0H	L0I
Geo		0.959 ± 0.016	
Rec		1.342 ± 0.005	
Trg	0.668 ± 0.005	2.921 ± 0.040	1.513 ± 0.013
PID	1.002 ± 0.0004	0.981 ± 0.001	0.996 ± 0.0005
MVA	0.810 ± 0.004	1.064 ± 0.010	0.978 ± 0.009
Tot	0.698 ± 0.013	3.924 ± 0.094	1.897 ± 0.040

Table 6.9: Summary of the rare to resonant efficiency double ratios $(\varepsilon_{ee}/\varepsilon_{J/\psi(ee)})/(\varepsilon_{\mu\mu}/\varepsilon_{J/\psi(\mu\mu)})$.

6.4.1 Geometric efficiency

In order to save disk space, the simulated samples contain only candidates for which all the daughter particles are inside the LHCb detector acceptance. This corresponds to the requirement that each of the final particles has a polar angle, θ, between 10 and 400 mrad, which is applied as filter at the generation level before passing the events to the full detector and reconstruction simulation. For this reason, the $\varepsilon^{\text{geom}}$ efficiency needs to be determined using dedicated generator level samples, where the filter has been removed. The results are shown in Figure 6.8.

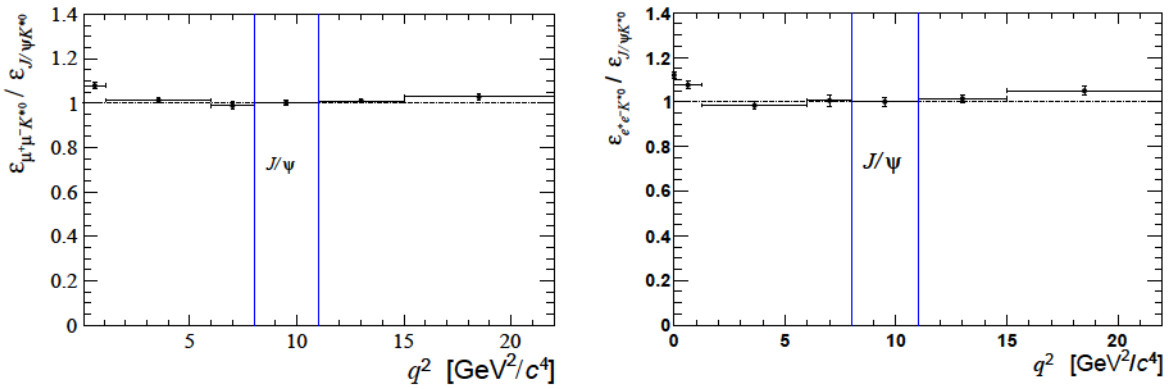


Figure 6.8: Ratio $\varepsilon_{K^{*0}\ell\ell}/\varepsilon_{K^{*0}J/\psi(\ell\ell)}$ of the geometric efficiencies for the rare and the resonant channels, in the case of muons (left) and electrons (right).

6.4.2 Reconstruction efficiency and bin migration

The $\varepsilon^{\text{reco|geom}}$ efficiency is defined as the number of reconstructed decays given that their decay products are inside the geometrical acceptance of the detector. This includes the probability for all tracks to be reconstructed, the decay to be stripped (Sec. 6.3.2) and pass the requirements that remove the backgrounds (Sec. 6.3.4). The PID requirements are removed from the stripping and accounted for using `PIDCalib`, as described in Sec. 6.4.3.

It has to be noted that, in the case of the electron channels, since the bremsstrahlung radiation of the electrons is not fully recovered, the q^2 resolution is worse than for muons. For this reason, events that are generated at a given q^2 might be reconstructed at a significantly different one. This effect is called bin migration. For example in the low q^2 region, due to the presence of the photon pole, there will be more events originally belonging to the lower edge of the bin that will not be reconstructed, than events that will come inside the region from

the central-q^2 bin. Figure 6.9 illustrates this phenomenon showing the correlation between the reconstructed and the generated q^2 for simulated $B^0 \to K^{*0} e^+ e^-$ events. Table 6.10 lists the net amounts of bin migration, M_{net}, in the J/ψ bin, which is defined as:

$$M_{\text{net}} = N(\text{in} \to \text{in}) + N(\text{out} \to \text{in}) - N(\text{in} \to \text{out}),$$

where $N(\text{in} \to \text{in})$ is the number of candidates that are generated and reconstructed inside the considered q^2 interval, $N(\text{out} \to \text{in})$ is the number of candidates that are generated outside the interval but reconstructed inside, and $N(\text{in} \to \text{out})$ is the number of candidates generated inside that fall outside.

Since the reconstruction efficiency is determined by comparing generated to reconstructed events, it already includes the bin migration. However, it is useful to single out this component to better asses the corresponding systematic uncertainty.

6.4.3 PID efficiency

The simulation is known to not reliably describe the PID variables and therefore the `PIDCalib` [7] package is used to determine the $\varepsilon^{\text{trig|PID}}$ efficiency. This package provides PID efficiency tables for each particle, based on high statistic data control samples, coming for example from charm decays. The efficiency in the tables are computed as function of momentum, pseudo-rapidity, number of tracks and final state particles. Applying these tables to our simulated events sample allows to obtain the correct PID efficiency.

6.4.4 Trigger efficiency

The trigger efficiency for the muon channels is determined using simulated events. A combination of simulation and data-driven methods is used instead for the electron modes, as described in sec:trigee. In particular, the electron efficiency of the trigger software stage, HLT, is measured from simulation, while the efficiency of the hardware stage, L0, is obtained from data driven techniques. As a cross-check, $J/\psi\,(\ell\ell)$ channels are used to test the data-simulation compatibility of $\varepsilon^{\text{trig|PID}}$ via the TISTOS approach, described in Sec. 6.4.4.2.

6.4.4.1 Electron triggers

Since the electron data are divided into three sub-samples, L0E, L0H and L0I, the $\varepsilon^{\text{trig|PID}}$ efficiency is separately measured for each trigger category. The HLT efficiency is determined using the simulation for all the three categories. Instead, the L0E and L0H efficiencies are measured using data. This is needed because the ageing of the calorimeters, on which the decision of these triggers relies on, is not well simulated. Tables of the efficiency depending on the p_T of the particles and the ECAL regions, obtained from data control samples via the TISTOS method and provided by the `L0CaloTool` tool, are used to calculate the following event probabilities:

$$P_{\texttt{L0Electron\_TOS}} = \varepsilon^{\text{L0Electron}}(e^+) + \varepsilon^{\text{L0Electron}}(e^-) - \varepsilon^{\text{L0Electron}}(e^+) \cdot \varepsilon^{\text{L0Electron}}(e^-),$$
$$P_{\texttt{L0Hadron\_TOS}} = \varepsilon^{\text{L0Hadron}}(K) + \varepsilon^{\text{L0Hadron}}(K) - \varepsilon^{\text{L0Hadron}}(\pi) \cdot \varepsilon^{\text{L0Hadron}}(\pi).$$

Event weights are then defined as:

- $w_{L0E} = P_{\texttt{L0Electron\_TOS}}$, corresponding to the probability that at least one of the electrons triggered;

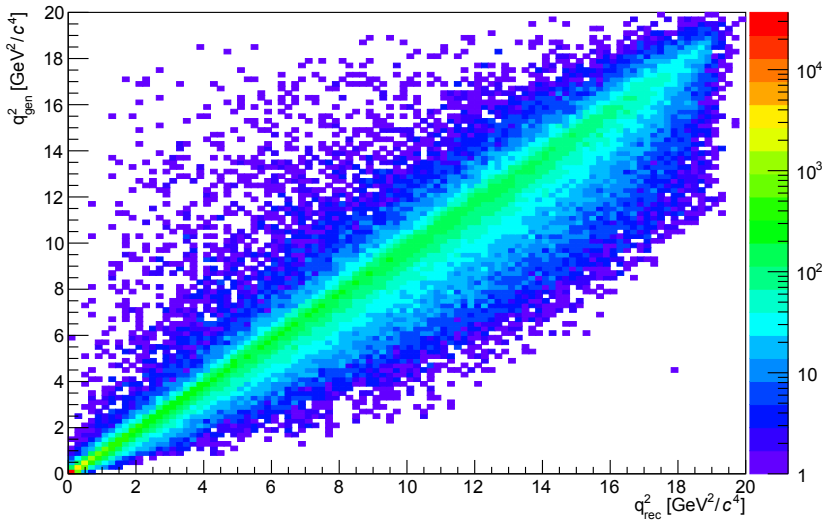


Figure 6.9: Generated versus reconstructed q^2 in simulated $B^0 \to K^{*0} e^+ e^-$ events.

Table 6.10: Bin migration in the J/ψ bin: positive (negative) values indicates "net in" ("net out").

Sample	J/ψ
$\mu\mu$	-0.0012 ± 0.0000
ee	-0.0360 ± 0.0002

- $w_{L0H} = P_{\texttt{L0Hadron\_TOS}} \cdot (1 - P_{\texttt{L0Electron\_TOS}})$, corresponding to the probability that at least one of the hadrons triggered, but none of the electrons;
- $w_{L0I} = (1 - P_{\texttt{L0Hadron\_TOS}}) \cdot (1 - P_{\texttt{L0Electron\_TOS}})$, corresponding to the probability that neither the hadrons nor the electrons triggered.

The L0E and L0H weights can be directly used to determine the trigger efficiency for the corresponding category as:

$$\varepsilon^{L0} = \frac{1}{N} \sum_{i}^{N} w_{L0},$$

This is not the case for the trigger efficiency of the L0I category. In this case, the efficiency is calculated counting how many events, on top of those passing the selection, satisfy the `L0Global_TIS` requirement in the simulation, after that the w_{L0I} data-driven weight is applied to "remove" the fraction of events that triggered L0E or L0H.

6.4.4.2 Crosscheck using the TISTOS method

The TISTOS [8] method is a powerful tool to determine the trigger efficiency. It can be used both on data and simulation, and therefore is useful to check for mis-modelling or biases. In this context it is used to test that the procedure described in the previous section provides the correct trigger efficiencies.

In the TISTOS method, two samples are defined:

- **TIS**: `L0Global_TIS && Hlt1Phys_TIS && Hlt2Phys_TIS`;
- **TOS**: `L0Global_TOS && Hlt1Phys_TOS && Hlt2Phys_TOS`.

Some events could belong both to TIS and TOS, and the efficiency for this to happen is defined as:

$$\epsilon_{\text{TIS|TOS}} = \frac{\text{N}_{\text{TISTOS}}}{\text{N}_{\text{TOS}}}$$

In the assumption that $\epsilon_{TIS} = \epsilon_{TIS|TOS}$, which is verified to be true in small bins of the kinematic phase space, one can evaluate a trigger efficiency relative to a generic selection as:

$$\epsilon_{Trig} = \frac{N_{Trig|Sel}}{N_{Sel}} = \frac{N_{Trig|Sel}}{N_{TIS|\text{Sel}}} \times \frac{N_{TIS|Sel}}{N_{\text{Sel}}} = \frac{N_{Trig|Sel}}{N_{TIS|Sel}} \times \epsilon_{TIS}$$

where all the quantities in the last term can then be determined directly from the sample, either a simulation or data collected by the detector.

This method is applied on the control sample $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$, both on data and simulation. For the data, as they contains a non negligible amount of background, the ${}_s\mathcal{P}lot$ technique is used to statistically remove it. For the simulation the L0 weights described in the previous Section 6.4.4.1 are applied.

Results are listed in Tab. 6.11 and shows a good compatibility between data and simulation for $J/\psi\,(\mu\mu)$, while there are significant deviations in $J/\psi\,(ee)$. The discrepancy is largest for the L0I category, for which indeed we know that the procedure described in Sec. 6.4.4.1 does not ensure a correct calibration.

Sample	**MC**	**Data**	f
$J/\psi\,(\mu\mu)$	0.797 ± 0.002	0.803 ± 0.004	1.0073
$J/\psi\,(ee)$ (L0E)	0.268 ± 0.002	0.255 ± 0.004	0.9536
$J/\psi\,(ee)$ (L0H)	0.028 ± 0.001	0.026 ± 0.002	0.9269
$J/\psi\,(ee)$ (L0I)	0.017 ± 0.001	0.011 ± 0.001	0.6760

Table 6.11: Trigger efficiencies determined using the TISTOS method on $B^0 \to K^{*0} J/\psi\,(\to \ell^+\ell^-)$ data and simulation.

For this reason, the $\varepsilon^{\text{trig|PID}}$ efficiency is corrected using the following factor:

$$f = 1 + \frac{\varepsilon^{\text{TISTOS, data}} - \varepsilon^{\text{TISTOS, MC}}}{\varepsilon^{\text{TISTOS, MC}}},$$

where the $\varepsilon^{\text{TISTOS}}$ is the trigger efficiency obtained for events triggering simultaneously TIS and TOS. For this correction to be effective on the charmonium channels as well as on the rare modes, the TISTOS efficiency must be reweighted for the difference in the kinematics between the two processes. This is done as function of the maximum p_{T} of the particles that fired L0: leptons for `L0Electron` and `L0Muon`, kaon and pion for `L0Hadron`, and all final state particles for `L0Global`. Results are shown in Fig. 6.10.

The ratios $\varepsilon^{\text{TISTOS}}_{\ell\ell} / \varepsilon^{\text{TISTOS}}_{J/\psi\,(\ell\ell)}$ obtained on data and simulation are found to be fully compatible, meaning that even though the TISTOS correction has an effect on the absolute efficiency, this becomes negligible on the ratio between the rare and resonant modes.

No correction is applied for the extraction of the R_{K^*} ratio, but this is accounted for in $R_{J/\psi}$.

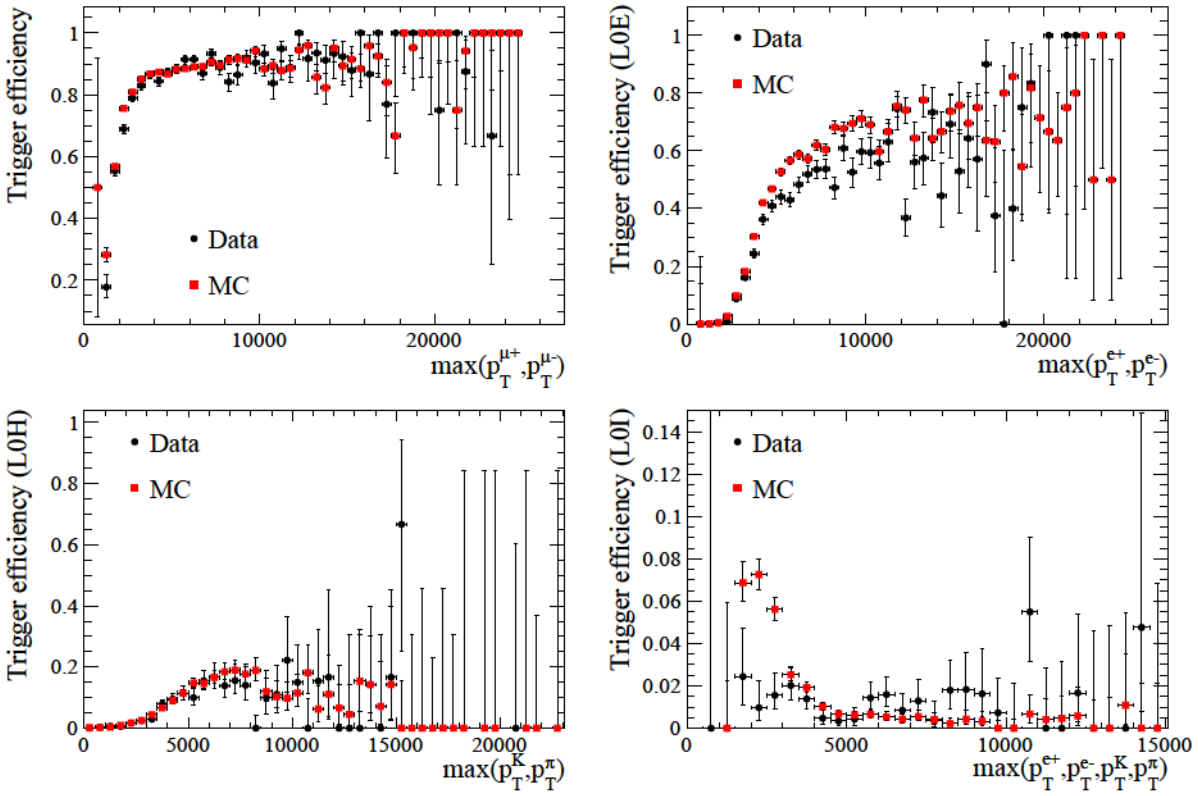


Figure 6.10: Trigger efficiency obtained using the TISTOS method on (top left) $B^0 \to K^{*0} J/\psi (\to \mu^+\mu^-)$ candidates as a function of the maximum p_T of the muons, and $B^0 \to K^{*0} J/\psi (\to e^+e^-)$ candidates as a function of the maximum p_T of the electrons (top left) for the L0E category, (bottom left) kaon and pion for L0H and (bottom right) all final particles for L0I.

6.4.5 Efficiency of the multivariate selection

The MVA efficiency is evaluated from fully weighted simulated samples and, for the electron channels, it is obtained separately for each trigger category, as the BDT cut is optimized independently depending on the category.

6.5 Mass fit

In order to extract the signal yields, a simultaneous unbinned maximum likelihood fit on the $m_{K\pi\ell^+\ell^-}$ invariant mass of the rare and resonant samples is performed. The yield of the rare channels is parameterised as a function of the $J/\psi(\ell\ell)$ yield as:

$$N_{\ell\ell} = N_{J/\psi(\ell\ell)} \cdot \frac{\varepsilon^{\ell\ell}}{\varepsilon^{J/\psi(\ell\ell)}} \cdot R_{\ell\ell},$$

where all the efficiency ratios between the rare and charmonium modes, $\varepsilon^{\ell\ell}/\varepsilon^{J/\psi(\ell\ell)}$, are given in Tab. 6.8. Consequently, $R_{\ell\ell}$ corresponds to the efficiency corrected ratio of the raw yields:

$$R_{\ell\ell} = \frac{N_{\ell\ell}/\varepsilon^{\ell\ell}}{N_{J/\psi(\ell\ell)}/\varepsilon^{J/\psi(\ell\ell)}}.$$

The R_{ee} and $R_{\mu\mu}$ ratios are then used to determine R_{K^*}, as described in Sec. 6.8. The following subsections describes the line shapes that are used to model all the components included in the fit.

6.5.1 Muon channels

For the resonant case, the 4-body invariant mass $m_{K\pi\mu\mu}$ is recalculated using a J/ψ mass constraint to the dilepton pair and a vertex constraint on the B meson. The effect of this kinematic fit is to improve the mass resolution by about a factor 2, which results in a more stable fit. This has also the advantage of pushing the mis-reconstructed background towards masses so low that it falls outside the fit region, which is chosen to be 5170–5850 MeV/c^2. As at this point there is no need to model the mis-reconstructed background in the fit, this also eliminates the systematic uncertainties associated with the knowledge of its shape.

The model chosen to describe the reconstructed B^0 invariant mass is a Double Crystal Ball (DCB) function with opposite-side tails, which has been already presented in Section 4.5.2, for both the rare and charmonium modes. Explicitly, the DCB function is parameterised as

$$\mathcal{P}_{\rm sig}(m|\vec{\lambda}) = f_{\rm core}\mathcal{P}_{\rm CB}(m|\mu,\sigma_1,\alpha_1,n_1) + (1-f_{\rm core})\mathcal{P}_{\rm CB}(m|\mu,\sigma_2,\alpha_2,n_2).$$

where $f_{\rm core}$ is the relative fraction of candidates falling in the first Crystal Ball function, σ_i the width of each CB, α_i and n_i are the parameters controlling the power law tail, and finally μ is the mean shared by the two Crystal Balls. In the first step, the parameters of the DCB are extracted using a fit on the invariant mass distribution of the $B^0 \to J/\psi K^{*0}$ simulated events samples, as shown on Fig 6.11.

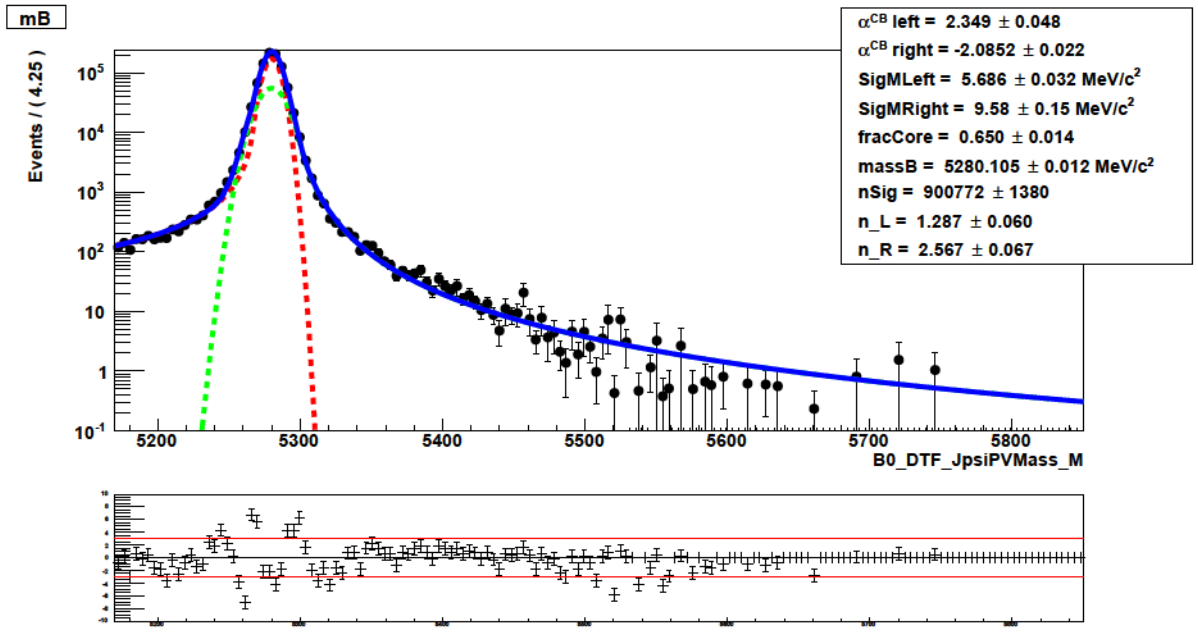


Figure 6.11: Fit to $m_{K\pi\mu\mu}$ mass distribution of $B^0 \to J/\psi K^{*0}$ simulated events.

In order to take into account possible data-MC differences in the invariant mass distribution, the μ is left free in the fit and the widths are allowed to scale differently for the charmonium and rare modes. All the other parameters are fixed to the values determined from the fit to simulation.

The combinatorial background is the only background component in the rare modes and it is described by an exponential function. In the resonant mode, a second signal component is included for the $B_s \to J/\psi \overline{K}^{*0}$ decay, expressed with the same signal parametrisation plus a shift of the mean by $\Delta m = m(B_s) - m(B^0)$. Finally, a $\Lambda_b^0 \to pKJ/\psi(\to \mu^+\mu^-)$ background component is modelled using simulated events to which the full $B^0 \to J/\psi K^{*0}$ selection is applied. This distribution has a broad shape under the signal peak and is parametrised from simulated samples using the `RooKeysPdf` class of the RooFit [9] package.

The invariant mass fit for the charmonium and the rare modes in the low q^2 region can be seen in Figure 6.12

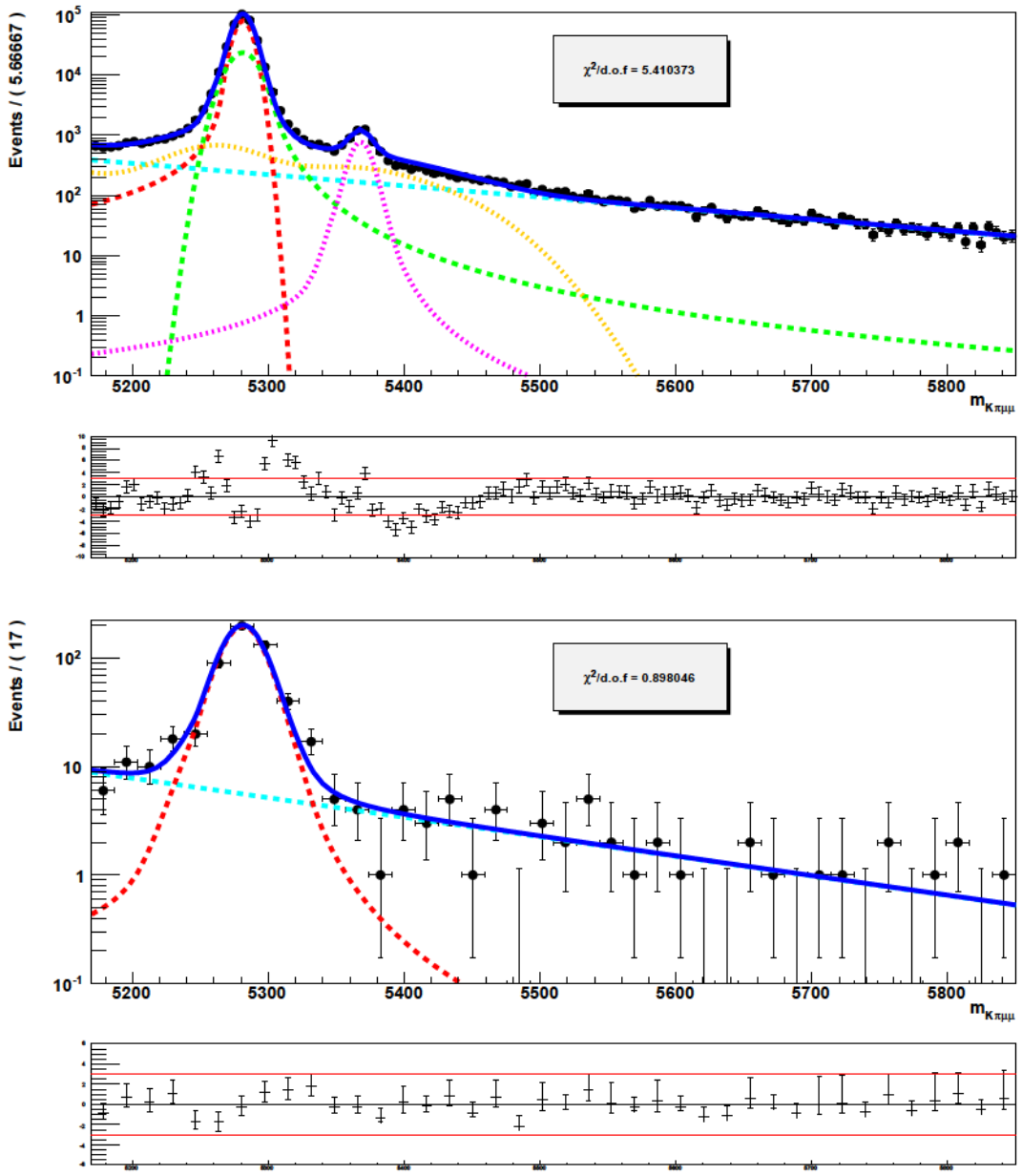


Figure 6.12: Fit to the $m_{K\pi\mu\mu}$ invariant mass of $B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-)$ candidates (top) and the rare mode in the low-q^2 region (bottom).

6.5.2 Electron channels

For the electron channels, a wider mass window, 4500 up to 5800 MeV/c^2, has been used to have a better control on the parameters which model the radiative tail and the mis-reconstructed backgrounds.

The reconstructed invariant mass of the B^0 depends on which L0 trigger line fired the event. Moreover, the shape strongly depends on how many bremsstrahlung photons were recovered. Three bremsstrahlung categories are considered:

- 0γ: events with no photon recovered;
- 1γ: events with one photon recovered from either of the electrons;
- 2γ: events with one photon recovered from each electron.

In the case where no bremsstrahlung photon has been found, the invariant mass distribution is modelled by a Crystal Ball function. For the 1γ and 2γ categories, the signal shape is parametrised by the sum of a CB function and a Gaussian function with all independent parameters. The Gaussian describes events in which a bremsstrahlung photon is added without being the proper one. All parameters are fixed using MC, and in order to take into account possible data-MC differences in the invariant mass distribution, the μ is left free in the fit and the widths are allowed to scale.

The results of the invariant mass fit on the $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ signal MC are shown in Figure 6.13.

A special selection is applied on $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ data to remove the partially reconstructed backgrounds and obtain a clean sample to be compared to MC. This special selection criteria is the request that the B mass computed with a J/ψ mass constrain is larger than 5175 MeV/c^2. Then the $m_{K\pi e^+e^-}$ distribution without the J/ψ mass constraint is fitted. An additional exponential component is added for the parametrisation of the combinatorial background. The results of the fit are shown in Figure 6.14.

For the fit of the invariant mass distribution in the rare mode, the $B^0 \to K^{*0} e^+e^-$ decay, a partially reconstructed background component has been added. The shape of the partially reconstructed events are extracted using a `RooKeysPdf` on a $B^+ \to \gamma K_1^+(1400)$ MC.

The results of the invariant mass fit on the $B^0 \to K^{*0} e^+e^-$ data are shown in Figure 6.15.

6.6 Systematic uncertainties

Since the R_{K^*} measurement is performed as a double ratio as shown in Eq. 6.1, many systematic uncertainties, that would have a significant impact on an absolute measurement, cancel.

6.7 Measurement of $R_{J/\psi}$

$R_{J/\psi}$ is the ratio between the branching ratio of the muon and electron resonant channels and has been used to cross-check the analysis procedure for two different regions: Low-q^2 and

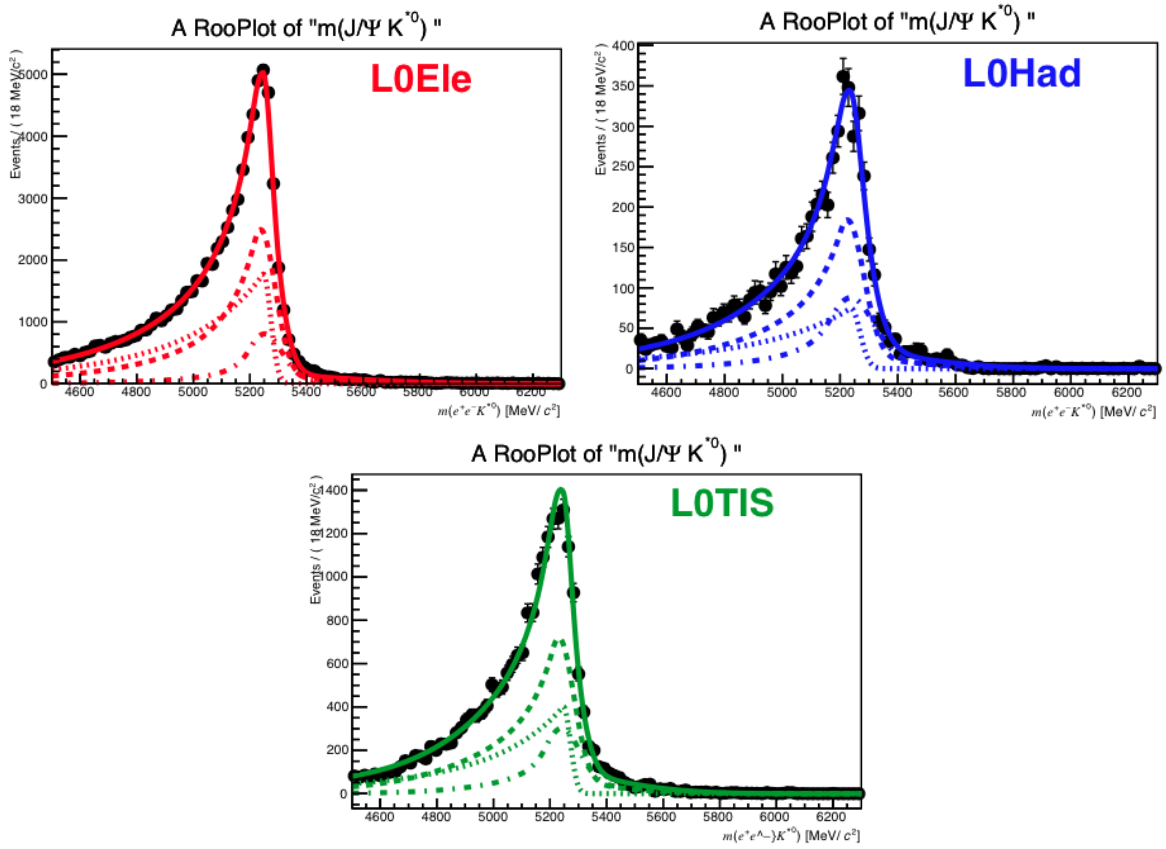


Figure 6.13: $m_{K\pi e^+e^-}$ invariant mass distribution for the $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ decay mode for the three trigger categories from MC. The dotted line is the signal PDF for the category without bremsstrahlung photon, the dashed one is for the category with one bremsstrahlung photon and finally the dashed-dotted one is for two or more bremsstrahlung photons category.

Central-q^2. It is determined as:

$$R_{J/\psi} = \frac{\mathcal{B}(B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-))}{\mathcal{B}(B^0 \to K^{*0} J/\psi\,(\to e^+e^-))} = \frac{\varepsilon_{J/\psi\,(\mu\mu)}}{\varepsilon_{J/\psi\,(ee)}} \cdot \frac{N_{B^0 \to K^{*0} J/\psi\,(\to e^+e^-)}}{N_{B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-)}}.$$

Here the trigger efficiencies needs to be corrected for data-simulation differences using the factors obtained in Sec. 6.4.4.2. The measured values of the $R_{J/\psi}$ ratios are reported in Tab. 6.12, where the error is only statistical. Note that not all systematic uncertainties that cancel on R_{K^*} fully cancel here. A reasonable agreement with unity is found, which validate the analysis in the different q^2 bins.

Trigger category	$R_{J/\psi}^{\text{Low-}q^2}$	$R_{J/\psi}^{\text{Central-}q^2}$
L0E	0.997 ± 0.007	1.028 ± 0.022
L0H	1.046 ± 0.025	0.986 ± 0.072
L0I	1.820 ± 0.026	0.973 ± 0.128

Table 6.12: Summary of the measured $\mathcal{B}(B^0 \to K^{*0} J/\psi\,(\to \mu^+\mu^-))/\mathcal{B}(B^0 \to K^{*0} J/\psi\,(\to e^+e^-))$ ratios.

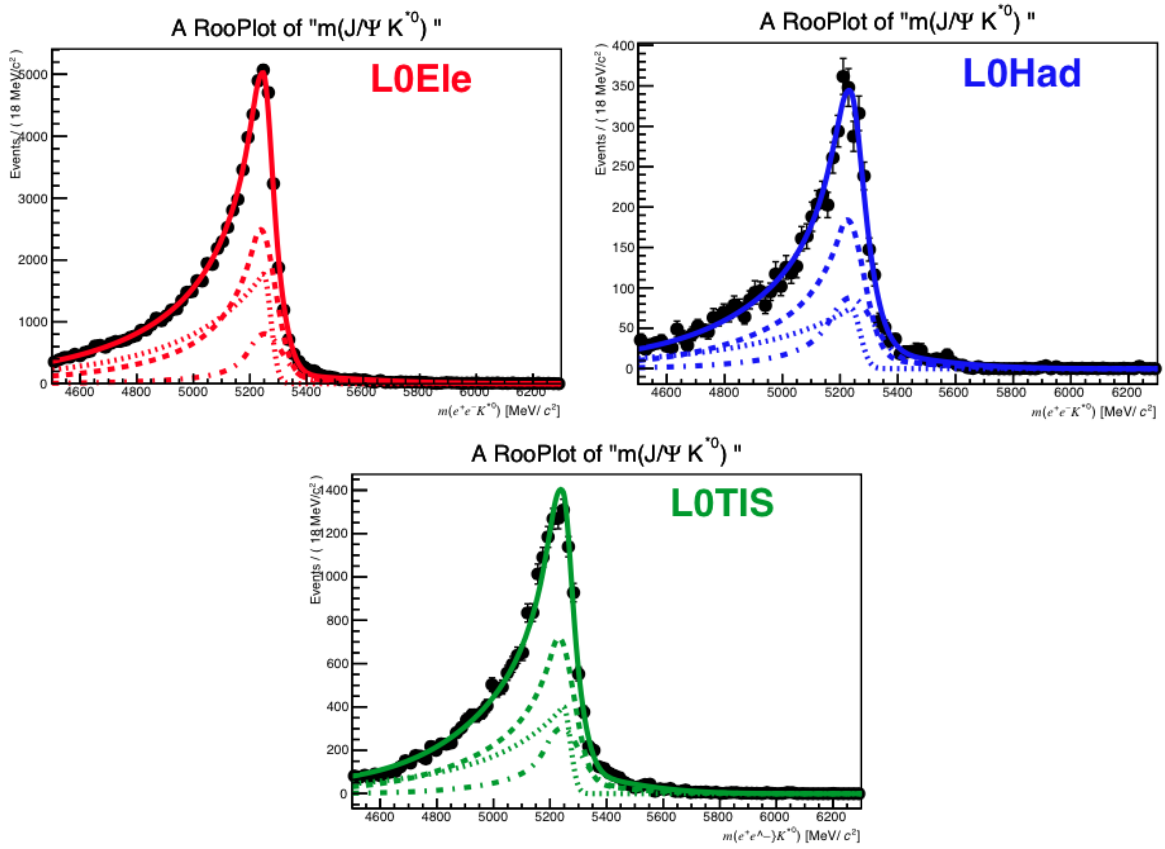


Figure 6.14: $m_{K\pi\mu\mu}$ invariant mass distribution for the $B^0 \to K^{*0} J/\psi (\to e^+e^-)$ decay mode for the three trigger categories in the data. The dotted line is the signal PDF for the category without bremsstrahlung photon, the dashed one is for the category with one bremsstrahlung photon and finally the dashed-dotted one is for two or more bremsstrahlung photons category.

6.8 Towards a result.

The R_{K^*} ratio will be calculated dividing the R_{ee} and $R_{\mu\mu}$ parameters determined from the mass fits in Sec. 6.5:

$$R_{K^*} = \frac{R_{\mu\mu}}{R_{ee}} = \frac{N_{\mu\mu}}{N_{J/\psi(\mu\mu)}} \cdot \frac{N_{J/\psi(ee)}}{N_{ee}} \cdot \frac{\varepsilon_{J/\psi(\mu\mu)}}{\varepsilon_{\mu\mu}} \cdot \frac{\varepsilon_{ee}}{\varepsilon_{J/\psi(ee)}}.$$

The advantage of using directly the electron ratio R_{ee} is that, since it is a shared parameter in the simultaneous fit to the three trigger categories, its value provides already a combination of the three samples.

The results are currently still blinded, as the analysis procedure is under review by other members of the LHCb collaboration, but they are expected to be published by the end of the year at latest.

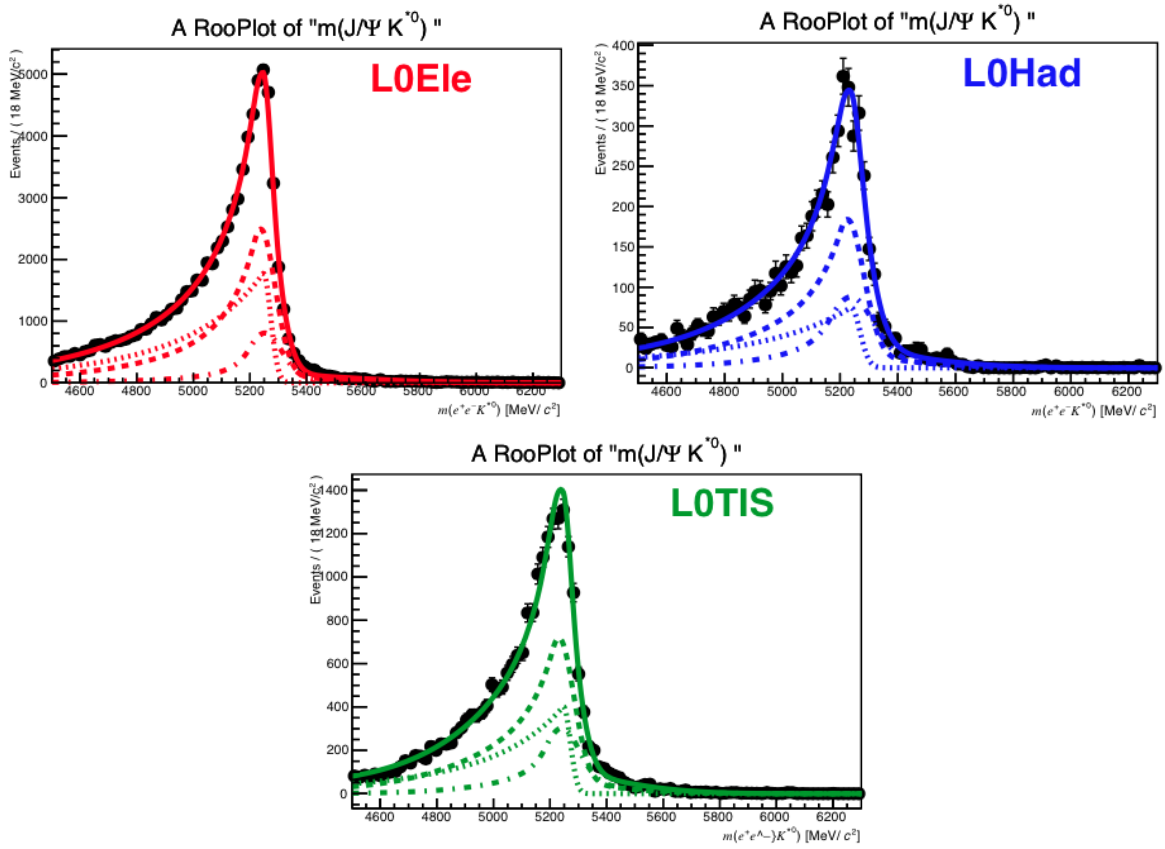


Figure 6.15: $m_{K\pi e^+e^-}$ invariant mass distribution for the $B^0 \to K^{*0}e^+e^-$ decay mode for the three trigger categories.

REFERENCES

[1] S. Bifani *et al.*, *Measurement of the ratio* $R_{K^*} = \mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)/\mathcal{B}(B^0 \to K^{*0}e^+e^-)$, LHCb-ANA-2015-016. Work in progress, at time of writing.

[2] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[3] T. Blake *et al.*, *Measurement of the full angular distribution of* $B^0 \to K^{*0}\mu^+\mu^-$ *decays using* $3\,\text{fb}^{-1}$ *of integrated luminosity*, LHCb-ANA-2013-097. Work in progress, at time of writing.

[4] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the* $B^0 \to K^{*0}e^+e^-$ *decay in the low-*q^2 *region*, JHEP **04** (2015) 024, arXiv:1501.0303.

[5] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using* $B^+ \to K^+\ell^+\ell^-$ *decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482.

[6] J. Lefrançois and M. H. Schune, *Measuring the photon polarization in b → s gamma using the B → K*e+e- decay channel*, Tech. Rep. LHCb-PUB-2009-008. CERN-LHCb-PUB-2009-008. LHCb-INT-2009-007, CERN, Geneva, Jun, 2009.

[7] LHCb collaboration, R. e. a. Aaij, *LHCb Detector Performance*, Int. J. Mod. Phys. A **30** (2014) 1530022. 82 p.

[8] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, Tech. Rep. LHCb-PUB-2014-039. CERN-LHCb-PUB-2014-039, CERN, Geneva, May, 2014.

[9] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116.

Conclusion

The angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay, performed on the whole run I data sample collected by the LHCb detector has been presented. This analysis, with respect to the previous one on $1fb^{-1}$ of data, benefits of an improved selection, based on a smaller number of discriminating variables but with a better global performances. For the first time, the full set of the angular observables has been measured through a maximum likelihood fit, allowing to provide to the theoreticians not only the measurements themselves, but also the correlations between the measurements. A larger number of smaller size q^2 bins has been used, thanks to the larger dataset available for this analysis. The results confirms the tendency saw already in the first analysis: a general good agreement of all the observables with the Standard Model predictions except for the P_5' variable, showing locally a discrepancy of 2.9 standard deviations in two bins.

Theoretical interpretations of these results, together with other related measurements, such as the differential branching ratios of the $B \to H\mu^+\mu^-$ or $B \to H\gamma$ decays (with H=K^*, ϕ, K, X_s, etc.) are already available. A result of such an approach[1] is shown in Figure. 6.16 for the New Physics Wilson coefficients $C_7^{\rm NP}$, $C_9^{\rm NP}$ and $C_{10}^{\rm NP}$. The best fit values are $C_9^{\rm NP} = -1.07$ with

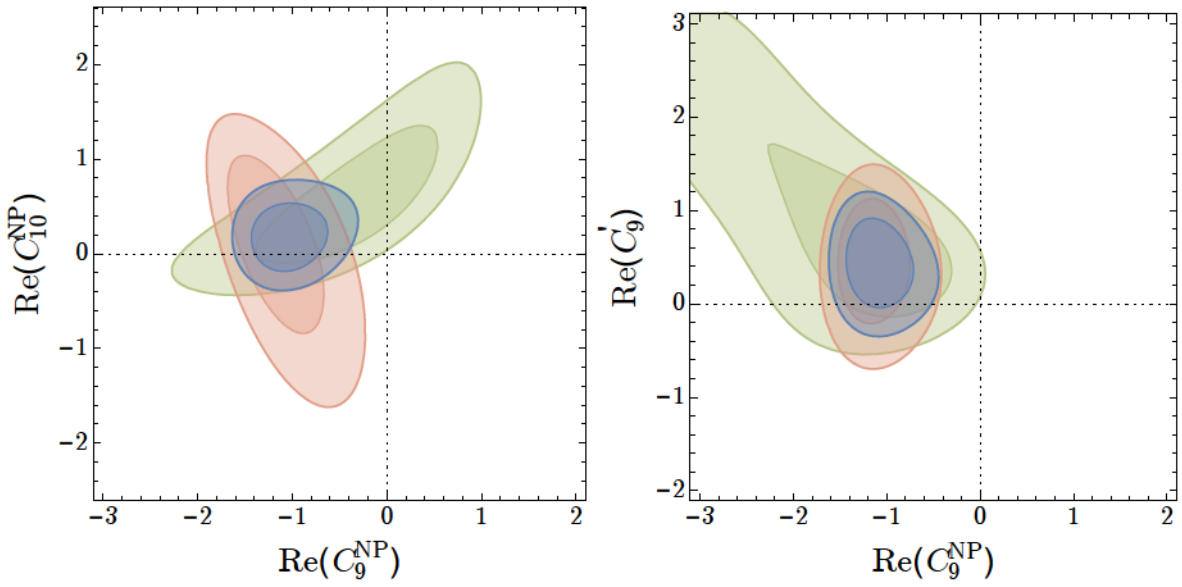


Figure 6.16: The blue area represents the best fit region from the global fit in the plane Re($C_9^{\rm NP}$)-Re($C_{10}^{\rm NP}$) plane (left) and in the Re($C_9^{\rm NP}$)-Re(C_9') plane (right). The red and green contours represent the allows region if only angular observables from $B^0 \to K^{*0}\mu^+\mu^-$ or only branching ratio data, respectively, is taken into account in the global fit.

[1] W. Altmannshofer and D. M. Straub, Implications of $b \to s$ measurements, in 50th Rencontres de Moriond on EW Interactions and Unified Theories La Thuile, Italy, March 14-21, 2015,

a p-value of 11.3% and $C_9^{\rm NP} = -C_{10}^{NP} = -0.53$ with a p-value =7.1%. These results could be explained by a new physics model with a new interaction mediated by a Z' particle with a mass of 7 TeV, but also by hadronic effects at unpredicted levels or by larger errors in the form factors computation. In addition, from an experimental point of view, the deviation from the Standard Model prediction is not yet at a significance level such to claim an evidence for new physics: a statistical fluctuation could still be responsible for this observation. As a consequence, more data are needed to further investigate this decay and confirm or not the result.
Some theoreticians relate the deviation in P_5' to the deviation of the LHCb R_K measurement from unity, supposing that the new interaction could violate the lepton flavour universality. A related analysis has been presented, the measurement of the R_{K^*} ratio, which is in itself a test of the lepton flavour universality and is also expected to add an important piece to the global puzzle of the $b \to s$ transitions. The analysis is still ongoing, but the first tests, and in particular the $R_{J/\psi}$ measurement, show that all its ingredients are well under control. The results are expected by the end of 2015.

In the future, the measurement of R_ϕ, the ratio $X = \frac{R_{K^*}}{R_K}$ proposed in[2] and of the ratios of angular observables will allow to further test the SM predictions using observables even more independent from the theoretical uncertainties and the experimental effects. The $B^0 \to K^{*0}\mu^+\mu^-$ decay will remain one of the main laboratory for the search of new physics in the years to come.

[2] G. Hiller and M. Schmaltz, Diagnosing lepton-nonuniversality in $b \to s\ell\ell$

Appendix A

Selection criteria

Candidate	Selection
B meson	IP $\chi^2 < 16$ (best PV)
B meson	$4600\,\text{MeV}/c^2 < M < 7000\,\text{MeV}/c^2$
B meson	DIRA angle $< 14\,\text{mrad}$
B meson	flight distance $\chi^2 > 121$
B meson	vertex $\chi^2/\text{ndf} < 8$
$\mu^+\mu^-$	$m(\mu^+\mu^-) < 7100\,\text{MeV}/c^2$
$\mu^+\mu^-$	vertex $\chi^2/\text{ndf} < 9$
K^{*0}	$m(K^+\pi^-) < 6200\,\text{MeV}/c^2$
K^{*0}	vertex $\chi^2/\text{ndf} < 9$
K^{*0}	flight distance $\chi^2 > 9$
tracks	ghost Prob < 0.4
tracks	min IP $\chi^2 > 9$
muon	`IsMuon`
muon	$\text{DLL}_{\mu\pi} > -3$
GEC	SPD Mult. < 600

Table A.1: Stripping selection criteria used in the 3fb^{-1} $B^0 \to K^{*0}\mu^+\mu^-$ analysis.

Candidates	Selection
Track	$0 < \theta < 400$ mrad
Track Pairs	$\theta_{pair} > 1$ mrad
$\mu^+\mu^-$	`IsMuon` True
K	`hasRich` True
K	$\text{DLL}_{K\pi}$ > -5
π	`hasRich` True
π	$\text{DLL}_{K\pi} < 25$
PV	$\lvert X - < X > \rvert < 5\,\text{mm}$
PV	$\lvert Y - < Y > \rvert < 5\,\text{mm}$
PV	$\lvert Z - < Z > \rvert < 200\,\text{mm}$

Table A.2: Pre-selection cuts applied to stripped candidates. In this table only: θ is the opening angle from the beam; and θ_{pair} is the opening angle between two track pairs. Variables $< X >$, $< Y >$, $< Z >$ denote the mean primary vertex position.

Appendix B

Simultaneous fit on invariant mass distribution

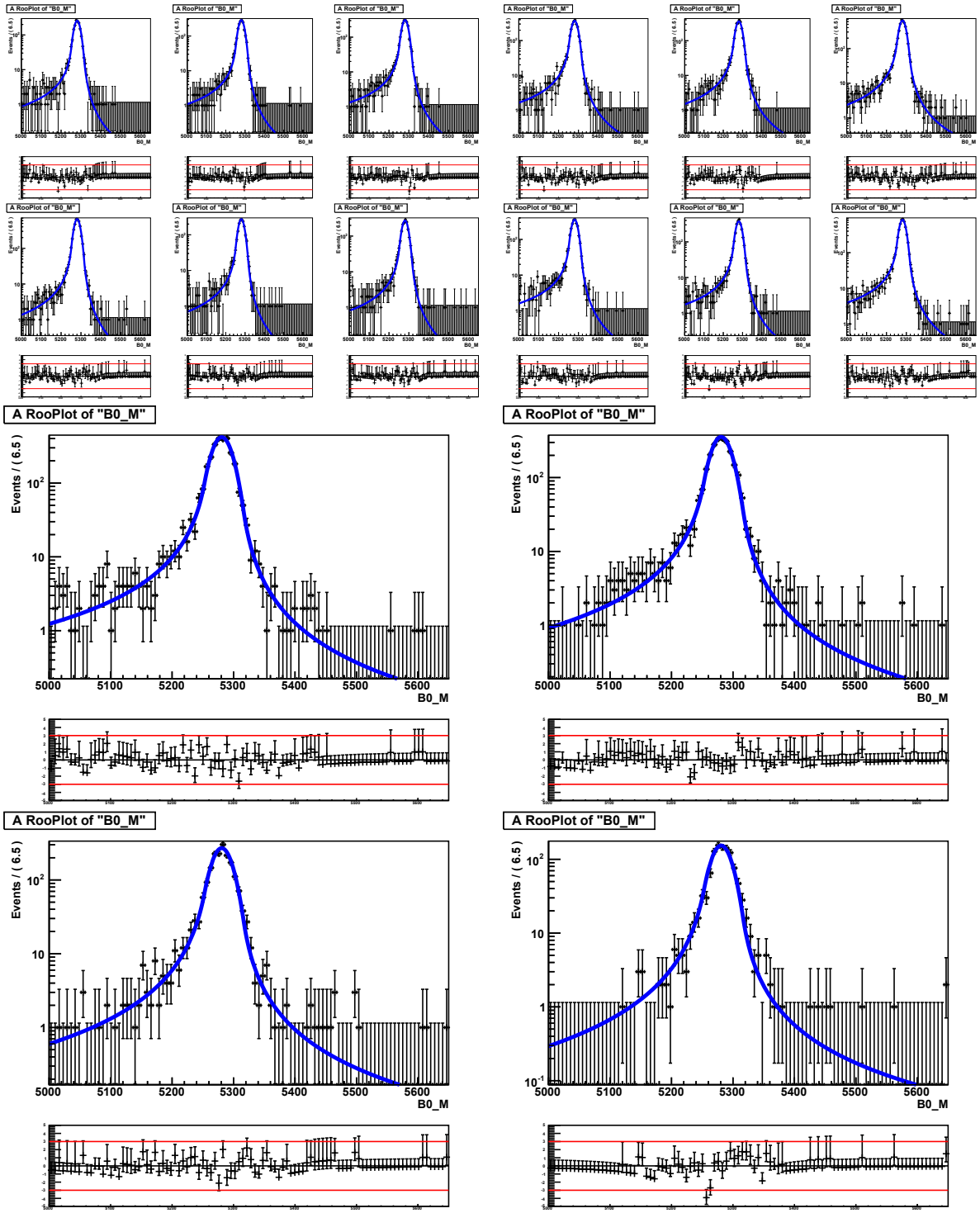
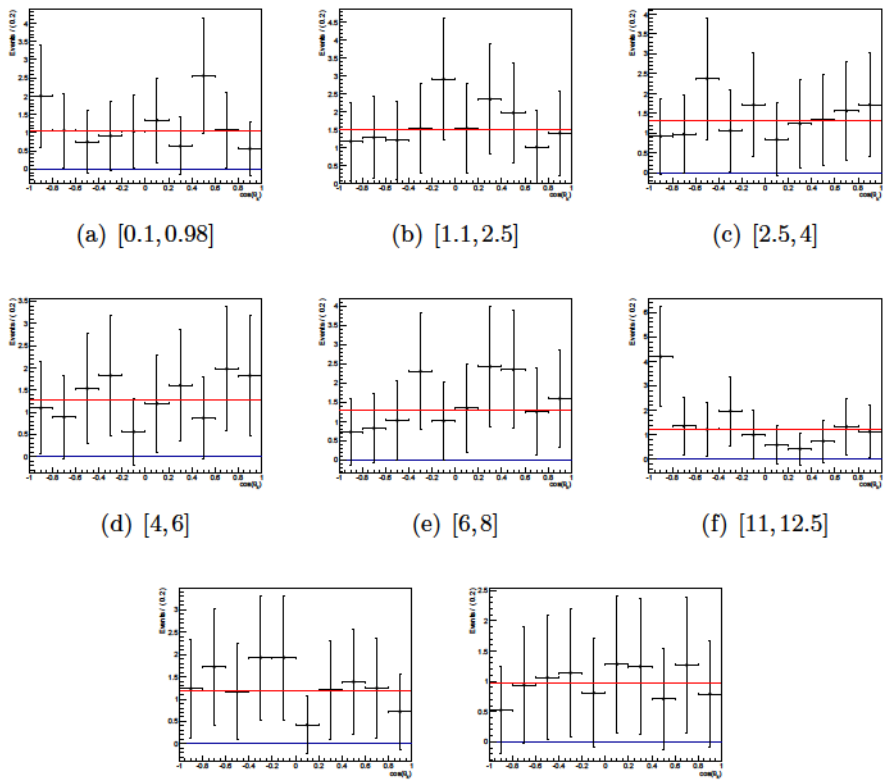


Figure B.1: Simultaneous fit on the K$\pi\mu\mu$ mass distribution fit of $B_d \to K^{*0}\mu\mu$ simulated events for 1 GeV/c^2 q^2 bins.

Appendix C

ABCD cross check

C.1 Angular correlation for the correction factor

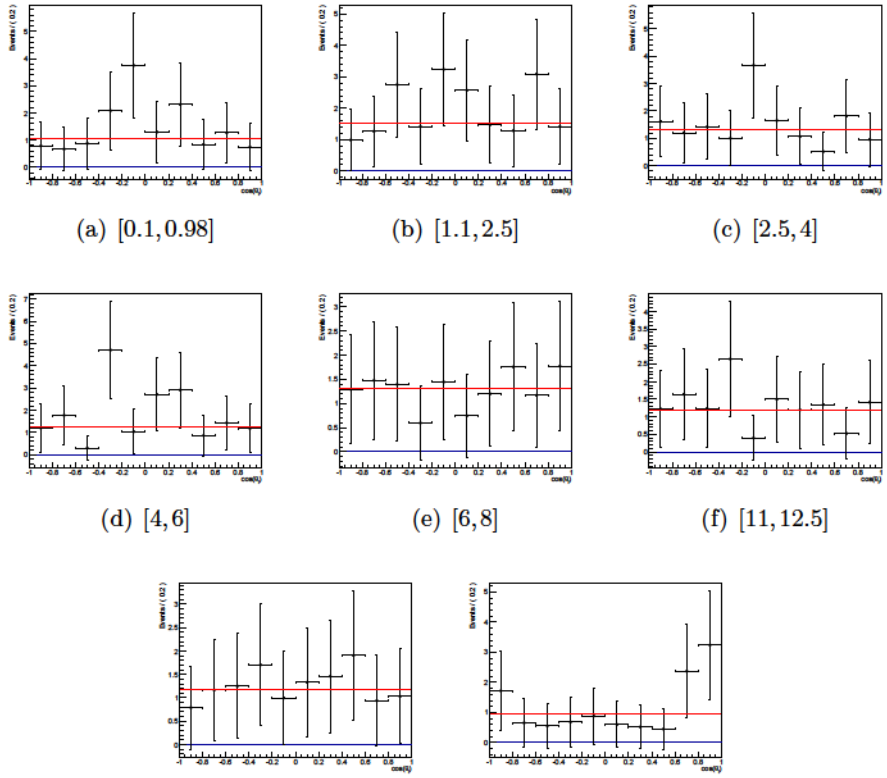


(a) $[0.1, 0.98]$ (b) $[1.1, 2.5]$ (c) $[2.5, 4]$

(d) $[4, 6]$ (e) $[6, 8]$ (f) $[11, 12.5]$

(g) $[15, 17]$ (h) $[17, 19]$

Figure C.1: Correction factor R for the $\cos(\theta_k)$ distribution for the 2 GeV^2/c^4 bin width.

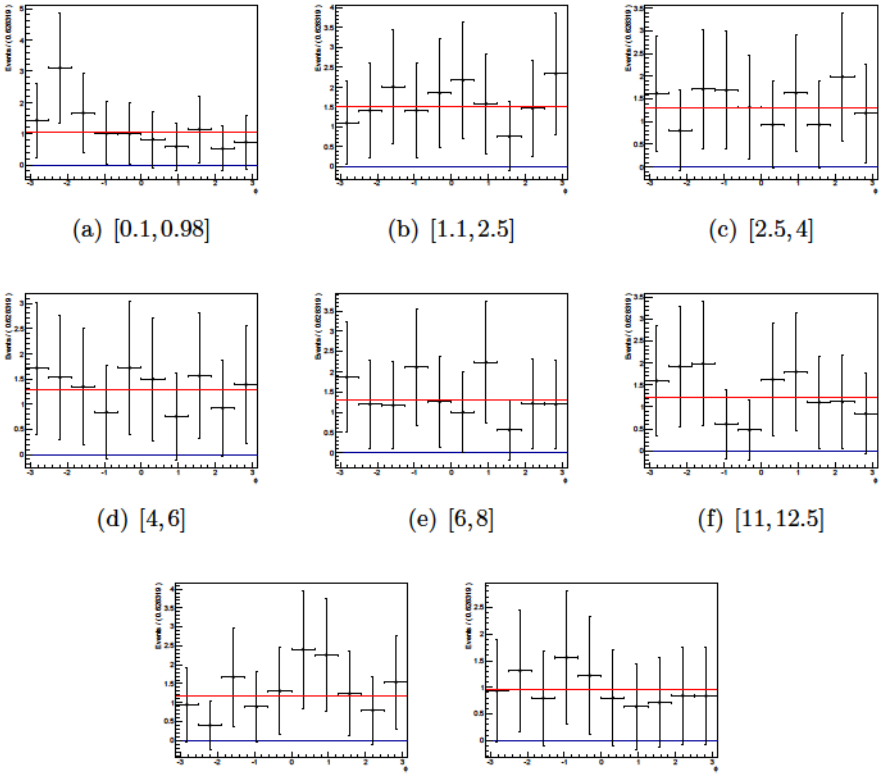


(a) $[0.1, 0.98]$ (b) $[1.1, 2.5]$ (c) $[2.5, 4]$

(d) $[4, 6]$ (e) $[6, 8]$ (f) $[11, 12.5]$

(g) $[15, 17]$ (h) $[17, 19]$

Figure C.2: Correction factor R for the $\cos(\theta_l)$ distribution for the 2 GeV^2/c^4 bin width.

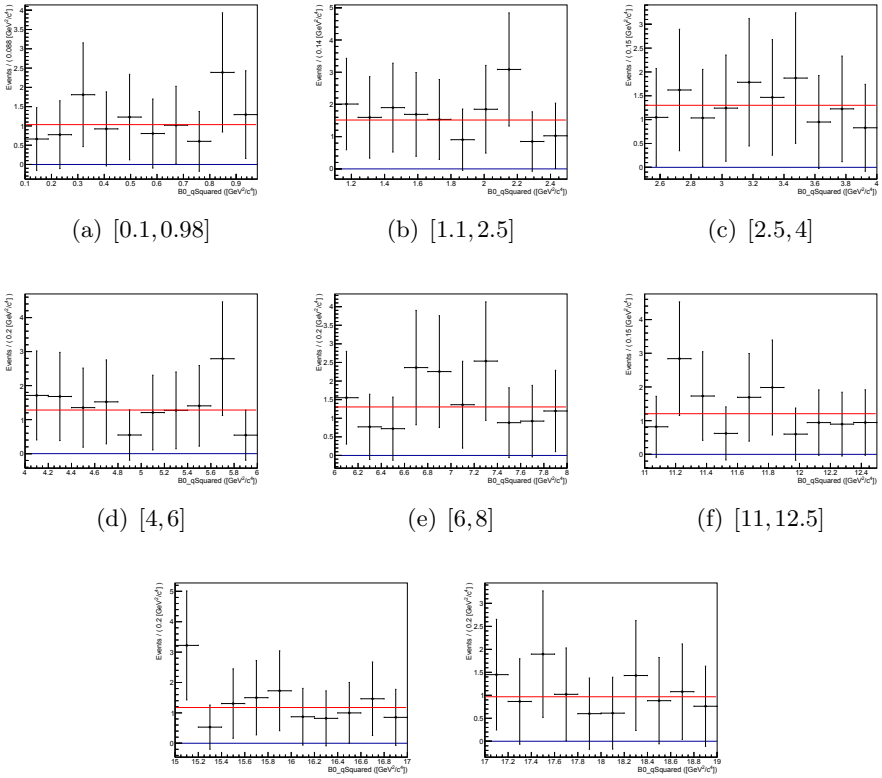


(a) $[0.1, 0.98]$ (b) $[1.1, 2.5]$ (c) $[2.5, 4]$

(d) $[4, 6]$ (e) $[6, 8]$ (f) $[11, 12.5]$

(g) $[15, 17]$ (h) $[17, 19]$

Figure C.3: Correction factor R for the ϕ distribution for the 2 GeV^2/c^4 bin width.



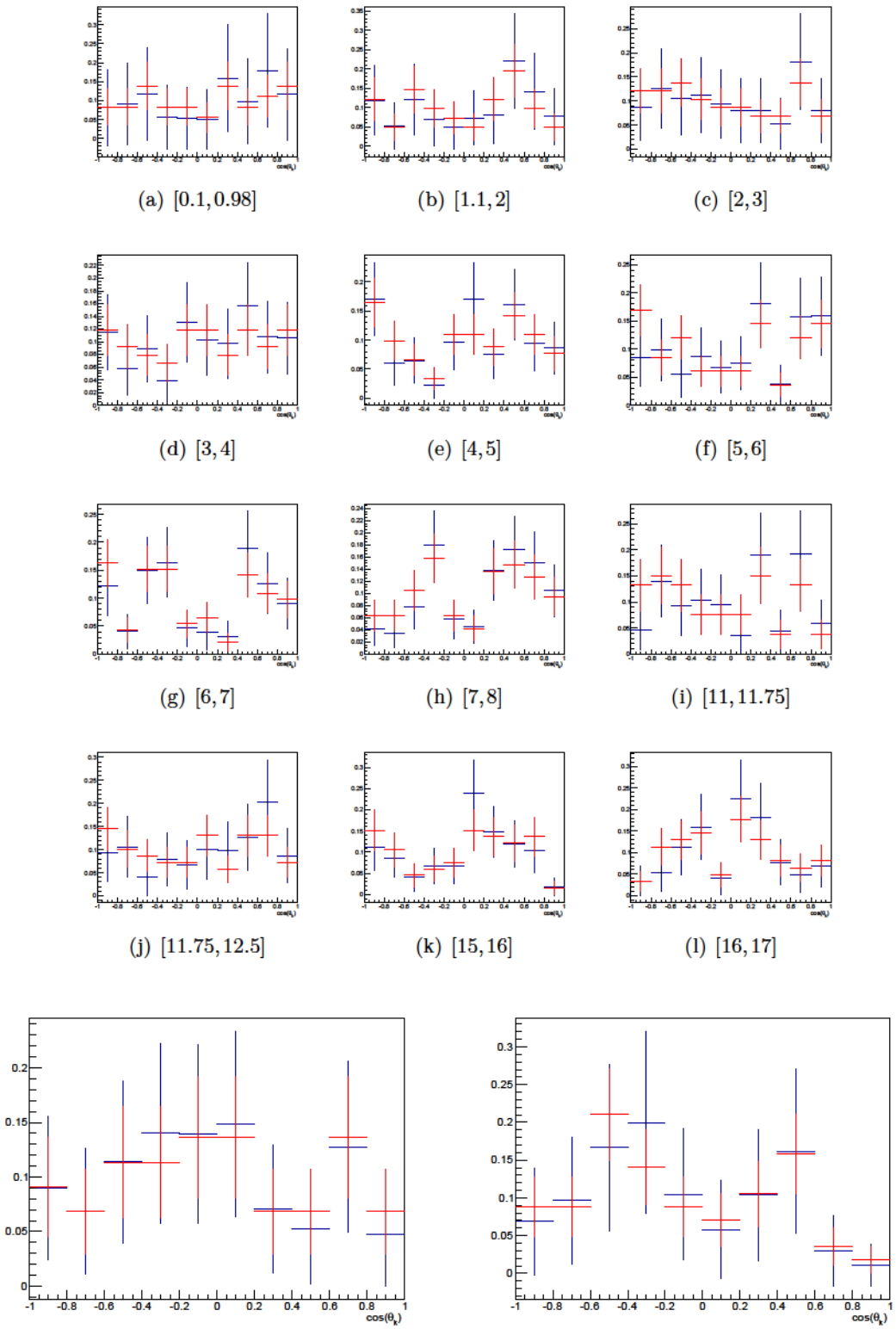
(a) $[0.1, 0.98]$ (b) $[1.1, 2.5]$ (c) $[2.5, 4]$

(d) $[4, 6]$ (e) $[6, 8]$ (f) $[11, 12.5]$

(g) $[15, 17]$ (h) $[17, 19]$

Figure C.4: Correction factor R for the q^2 distribution for the 2 GeV$^2/c^4$ bin width.

C.2 Comparison with upper mass sideband



(a) $[0.1, 0.98]$ (b) $[1.1, 2]$ (c) $[2, 3]$

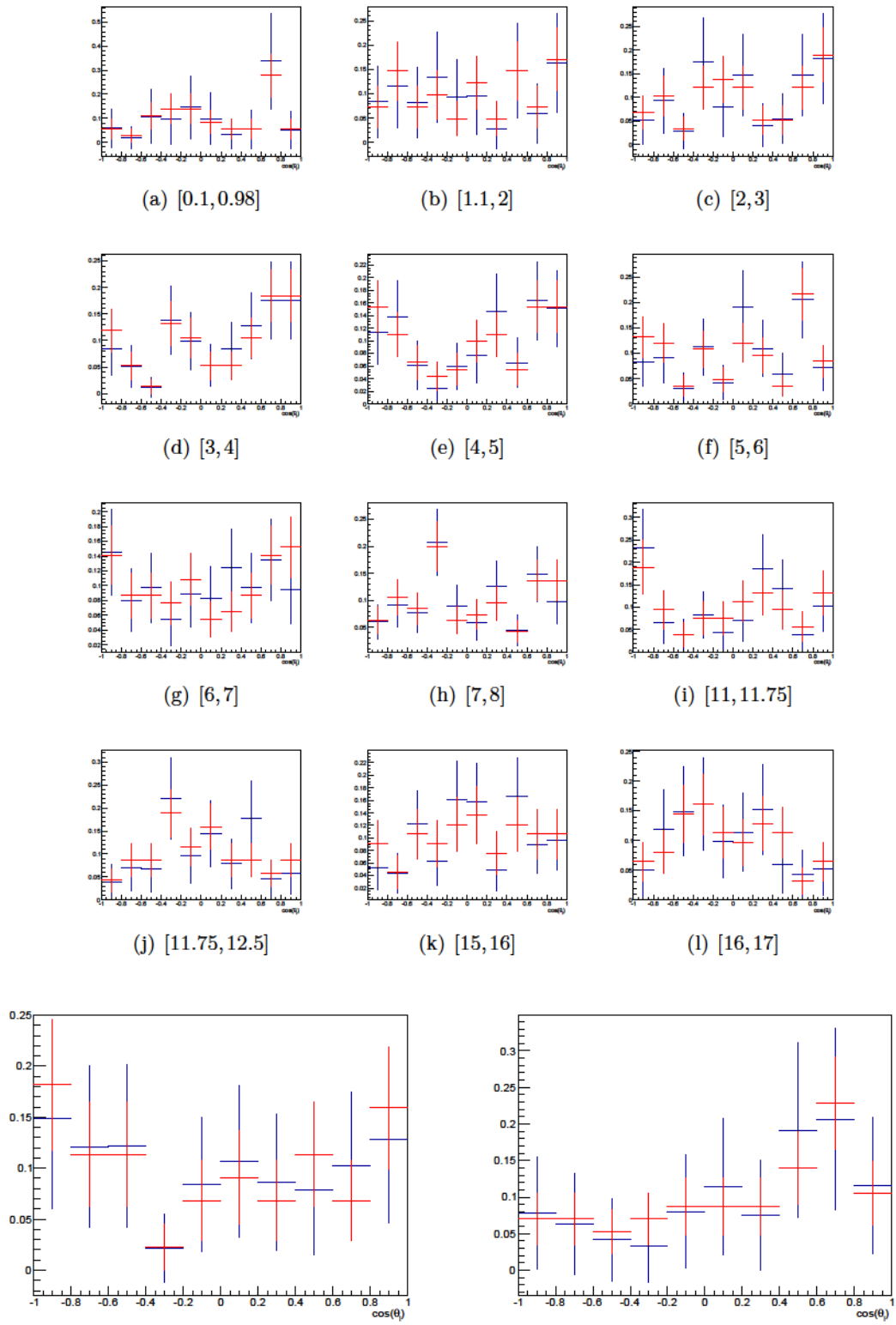
(d) $[3, 4]$ (e) $[4, 5]$ (f) $[5, 6]$

(g) $[6, 7]$ (h) $[7, 8]$ (i) $[11, 11.75]$

(j) $[11.75, 12.5]$ (k) $[15, 16]$ (l) $[16, 17]$

(m) $[17, 18]$ (n) $[18, 19]$

Figure C.5: Comparison of the $\cos(\theta_k)$ distribution for the ABCD method (blue) and the upper mass sideband (red) in the 1 GeV$^2/c^4$ binning scheme.



(a) $[0.1, 0.98]$ (b) $[1.1, 2]$ (c) $[2, 3]$

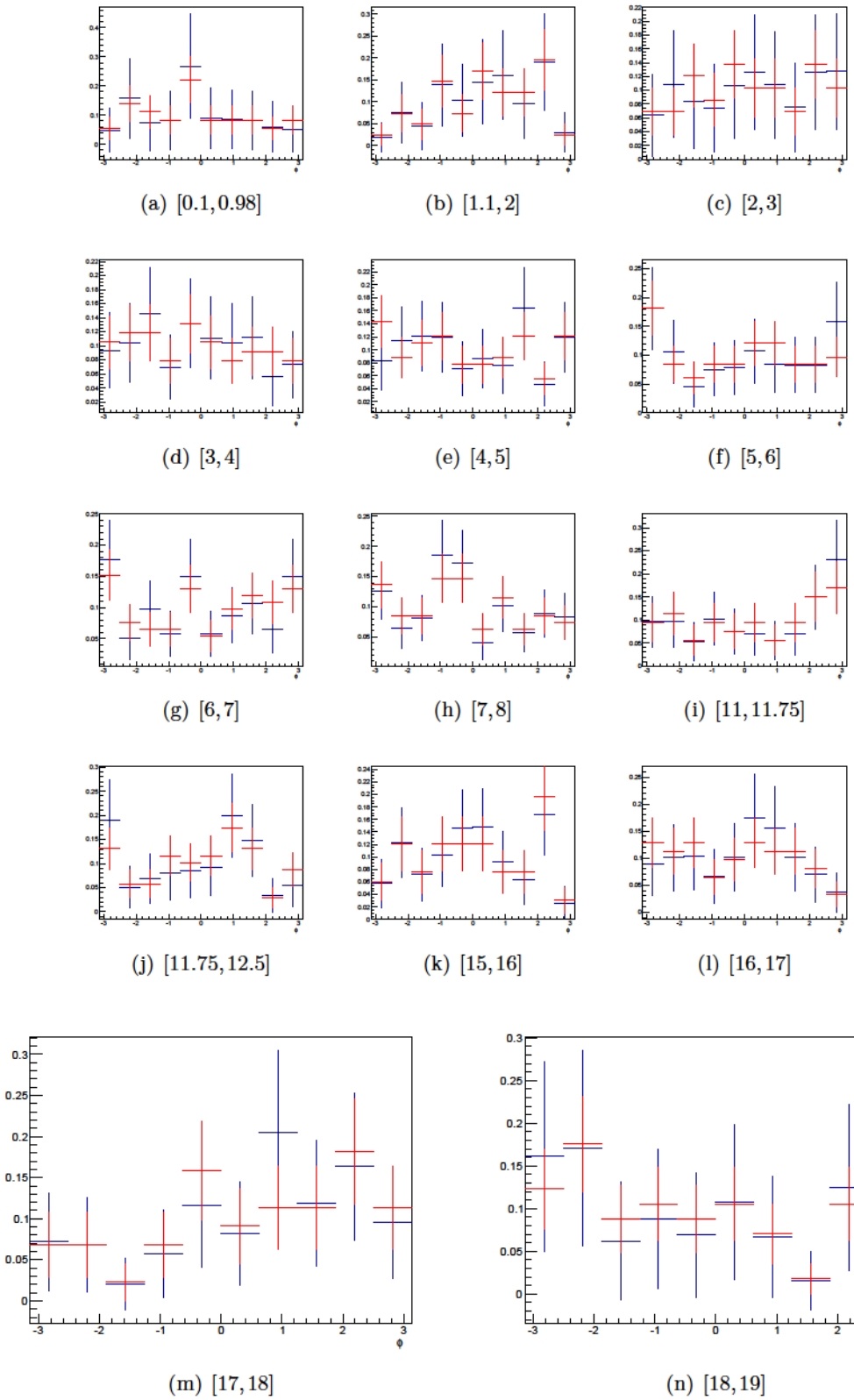
(d) $[3, 4]$ (e) $[4, 5]$ (f) $[5, 6]$

(g) $[6, 7]$ (h) $[7, 8]$ (i) $[11, 11.75]$

(j) $[11.75, 12.5]$ (k) $[15, 16]$ (l) $[16, 17]$

(m) $[17, 18]$ (n) $[18, 19]$

Figure C.6: Comparison of the $\cos(\theta_l)$ distribution for the ABCD method (blue) and the upper mass sideband (red) in the 1 GeV$^2/c^4$ binning scheme.



(a) [0.1, 0.98] (b) [1.1, 2] (c) [2, 3]

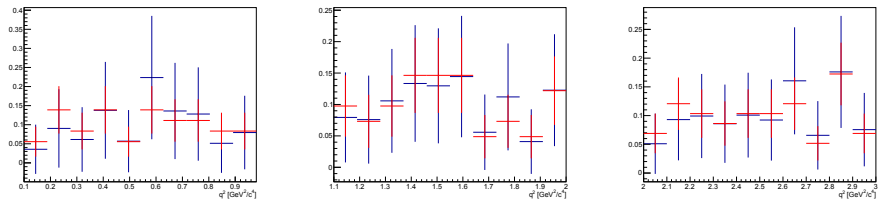
(d) [3, 4] (e) [4, 5] (f) [5, 6]

(g) [6, 7] (h) [7, 8] (i) [11, 11.75]

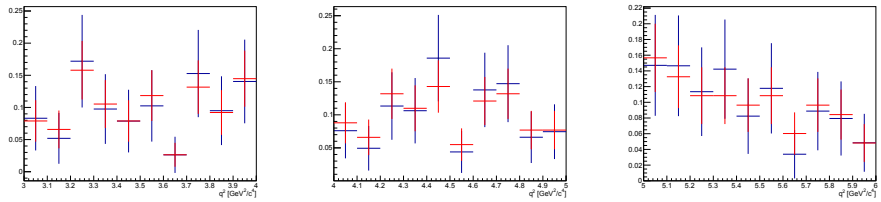
(j) [11.75, 12.5] (k) [15, 16] (l) [16, 17]

(m) [17, 18] (n) [18, 19]

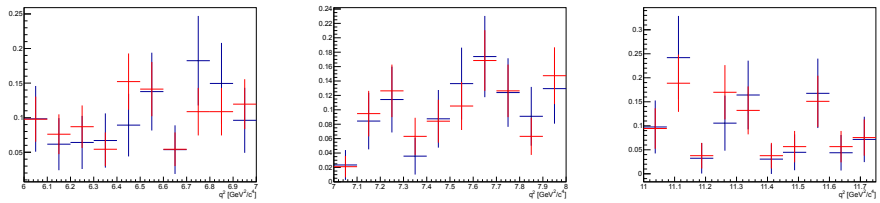
Figure C.7: Comparison of the ϕ distribution for the ABCD method (blue) and the upper mass sideband (red) in the 1 GeV$^2/c^4$ binning scheme.



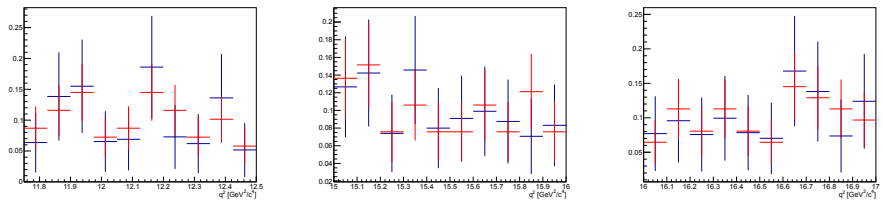
(a) $[0.1, 0.98]$ (b) $[1.1, 2]$ (c) $[2, 3]$



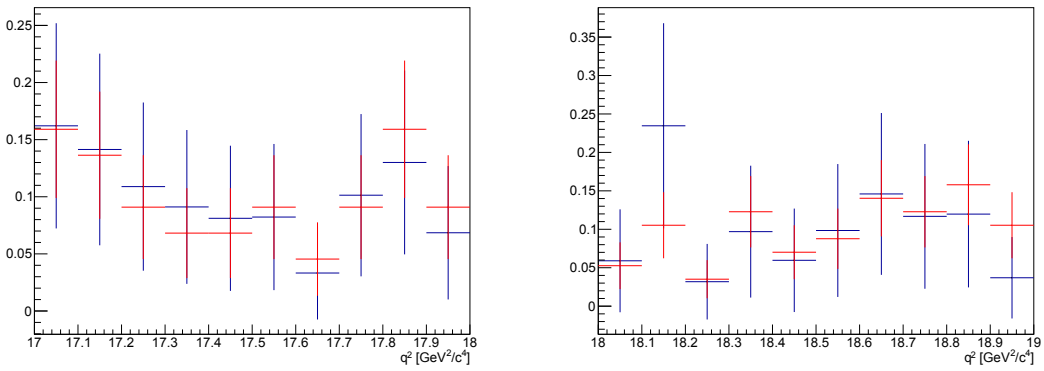
(d) $[3, 4]$ (e) $[4, 5]$ (f) $[5, 6]$



(g) $[6, 7]$ (h) $[7, 8]$ (i) $[11, 11.75]$



(j) $[11.75, 12.5]$ (k) $[15, 16]$ (l) $[16, 17]$



(m) $[17, 18]$ (n) $[18, 19]$

Figure C.8: Comparison of the q^2 distribution for the ABCD method (blue) and the upper mass sideband (red) in the 1 GeV$^2/c^4$ binning scheme.

Appendix D

Systematics test

D.1 Statistical uncertainty of the four-dimensional acceptance

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0029	F_L	0.0018	F_L	0.0013	F_L	0.0012	F_L	0.0013
S_3	0.0038	S_3	0.0015	S_3	0.0012	S_3	0.0012	S_3	0.0013
S_4	0.0040	S_4	0.0025	S_4	0.0018	S_4	0.0015	S_4	0.0012
S_5	0.0045	S_5	0.0026	S_5	0.0022	S_5	0.0018	S_5	0.0018
A_{FB}	0.0038	A_{FB}	0.0013	A_{FB}	0.0009	A_{FB}	0.0008	A_{FB}	0.0009
S_7	0.0003	S_7	0.0002	S_7	0.0001	S_7	0.0001	S_7	0.0001
S_8	0.0001	S_8	0.0001	S_8	0.0001	S_8	0.0000	S_8	0.0000
S_9	0.0000	S_9	0.0001	S_9	0.0000	S_9	0.0000	S_9	0.0000

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0025	F_L	0.0044	F_L	0.0017	F_L	0.0012	F_L	0.0029
S_3	0.0030	S_3	0.0067	S_3	0.0017	S_3	0.0011	S_3	0.0039
S_4	0.0021	S_4	0.0037	S_4	0.0012	S_4	0.0016	S_4	0.0023
S_5	0.0028	S_5	0.0046	S_5	0.0019	S_5	0.0018	S_5	0.0031
A_{FB}	0.0020	A_{FB}	0.0037	A_{FB}	0.0012	A_{FB}	0.0007	A_{FB}	0.0022
S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0001	S_7	0.0001
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0001	S_8	0.0001
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

Table D.1: Systematic uncertainties due to the statistical uncertainty on the four-dimensional acceptance. Ranges of q^2 bins are given in GeV^2/c^4.

D.2 Differences between data and simulation

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0009	F_L	0.0028	F_L	0.0027	F_L	0.0032	F_L	0.0039
S_3	0.0002	S_3	0.0000	S_3	0.0001	S_3	0.0001	S_3	0.0001
S_4	0.0003	S_4	0.0002	S_4	0.0005	S_4	0.0004	S_4	0.0001
S_5	0.0004	S_5	0.0005	S_5	0.0001	S_5	0.0005	S_5	0.0005
A_{FB}	0.0003	A_{FB}	0.0014	A_{FB}	0.0009	A_{FB}	0.0005	A_{FB}	0.0016
S_7	0.0000	S_7	0.0000	S_7	0.0001	S_7	0.0000	S_7	0.0000
S_8	0.0001	S_8	0.0000	S_8	0.0001	S_8	0.0000	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0032	F_L	0.0024	F_L	0.0046	F_L	0.0032	F_L	0.0029
S_3	0.0001	S_3	0.0005	S_3	0.0005	S_3	0.0000	S_3	0.0002
S_4	0.0001	S_4	0.0005	S_4	0.0006	S_4	0.0002	S_4	0.0003
S_5	0.0005	S_5	0.0003	S_5	0.0009	S_5	0.0001	S_5	0.0004
A_{FB}	0.0017	A_{FB}	0.0011	A_{FB}	0.0028	A_{FB}	0.0005	A_{FB}	0.0015
S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

Table D.2: Systematic uncertainties from neglecting the explicit reweighting of the B^0 p_{T}. Ranges of q^2 bins are given in GeV^2/c^4.

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0019	F_L	0.0004	F_L	0.0005	F_L	0.0004	F_L	0.0004
S_3	0.0001	S_3	0.0001	S_3	0.0001	S_3	0.0001	S_3	0.0000
S_4	0.0019	S_4	0.0001	S_4	0.0003	S_4	0.0002	S_4	0.0001
S_5	0.0004	S_5	0.0006	S_5	0.0004	S_5	0.0000	S_5	0.0001
A_{FB}	0.0019	A_{FB}	0.0003	A_{FB}	0.0001	A_{FB}	0.0001	A_{FB}	0.0002
S_7	0.0002	S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000
S_9	0.0001	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0010	F_L	0.0013	F_L	0.0010	F_L	0.0005	F_L	0.0011
S_3	0.0002	S_3	0.0007	S_3	0.0000	S_3	0.0001	S_3	0.0004
S_4	0.0003	S_4	0.0004	S_4	0.0001	S_4	0.0003	S_4	0.0003
S_5	0.0004	S_5	0.0002	S_5	0.0001	S_5	0.0003	S_5	0.0003
A_{FB}	0.0004	A_{FB}	0.0002	A_{FB}	0.0006	A_{FB}	0.0000	A_{FB}	0.0003
S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

Table D.3: Systematic uncertainties form neglecting the explicit reweighting of the B^0 vertex χ^2. Ranges of q^2 bins are given in GeV^2/c^4.

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0010	F_L	0.0002	F_L	0.0003	F_L	0.0004	F_L	0.0007
S_3	0.0000	S_3	0.0002	S_3	0.0002	S_3	0.0000	S_3	0.0001
S_4	0.0005	S_4	0.0010	S_4	0.0010	S_4	0.0008	S_4	0.0006
S_5	0.0003	S_5	0.0003	S_5	0.0001	S_5	0.0001	S_5	0.0001
A_{FB}	0.0022	A_{FB}	0.0000	A_{FB}	0.0005	A_{FB}	0.0006	A_{FB}	0.0006
S_7	0.0001	S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000
S_8	0.0002	S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0010	F_L	0.0015	F_L	0.0003	F_L	0.0000	F_L	0.0003
S_3	0.0001	S_3	0.0001	S_3	0.0007	S_3	0.0002	S_3	0.0002
S_4	0.0002	S_4	0.0009	S_4	0.0003	S_4	0.0011	S_4	0.0003
S_5	0.0004	S_5	0.0001	S_5	0.0005	S_5	0.0003	S_5	0.0002
A_{FB}	0.0006	A_{FB}	0.0016	A_{FB}	0.0002	A_{FB}	0.0004	A_{FB}	0.0005
S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

Table D.4: Systematic uncertainties form neglecting the explicit reweighting of the track multiplicity. Ranges of q^2 bins are given in GeV^2/c^4.

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0139	F_L	0.0149	F_L	0.0118	F_L	0.0126	F_L	0.0130
S_3	0.0010	S_3	0.0002	S_3	0.0005	S_3	0.0011	S_3	0.0014
S_4	0.0005	S_4	0.0017	S_4	0.0007	S_4	0.0021	S_4	0.0020
S_5	0.0030	S_5	0.0006	S_5	0.0017	S_5	0.0020	S_5	0.0010
A_{FB}	0.0003	A_{FB}	0.0077	A_{FB}	0.0043	A_{FB}	0.0020	A_{FB}	0.0058
S_7	0.0003	S_7	0.0001	S_7	0.0003	S_7	0.0002	S_7	0.0001
S_8	0.0000	S_8	0.0001	S_8	0.0002	S_8	0.0002	S_8	0.0001
S_9	0.0001	S_9	0.0001	S_9	0.0002	S_9	0.0002	S_9	0.0001

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0059	F_L	0.0021	F_L	0.0108	F_L	0.0139	F_L	0.0040
S_3	0.0020	S_3	0.0019	S_3	0.0015	S_3	0.0004	S_3	0.0021
S_4	0.0006	S_4	0.0010	S_4	0.0002	S_4	0.0001	S_4	0.0009
S_5	0.0018	S_5	0.0011	S_5	0.0016	S_5	0.0011	S_5	0.0015
A_{FB}	0.0031	A_{FB}	0.0005	A_{FB}	0.0066	A_{FB}	0.0023	A_{FB}	0.0017
S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0002	S_7	0.0000
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0001	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0001	S_9	0.0000

Table D.5: Systematic uncertainties from reweighting depending on pion p and p_{T}. Ranges of q^2 bins are given in GeV^2/c^4.

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0035	F_L	0.0006	F_L	0.0001	F_L	0.0004	F_L	0.0008
S_3	0.0010	S_3	0.0002	S_3	0.0001	S_3	0.0001	S_3	0.0002
S_4	0.0003	S_4	0.0006	S_4	0.0010	S_4	0.0010	S_4	0.0008
S_5	0.0010	S_5	0.0004	S_5	0.0000	S_5	0.0004	S_5	0.0004
A_{FB}	0.0008	A_{FB}	0.0004	A_{FB}	0.0001	A_{FB}	0.0001	A_{FB}	0.0001
S_7	0.0002	S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000
S_8	0.0001	S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000
S_9	0.0001	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0008	F_L	0.0002	F_L	0.0012	F_L	0.0000	F_L	0.0003
S_3	0.0002	S_3	0.0006	S_3	0.0001	S_3	0.0000	S_3	0.0004
S_4	0.0000	S_4	0.0001	S_4	0.0002	S_4	0.0013	S_4	0.0000
S_5	0.0000	S_5	0.0002	S_5	0.0002	S_5	0.0000	S_5	0.0001
A_{FB}	0.0005	A_{FB}	0.0001	A_{FB}	0.0008	A_{FB}	0.0001	A_{FB}	0.0003
S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

Table D.6: Systematic uncertainties from reweighting depending on kaon p and p_{T}. Ranges of q^2 bins are given in GeV^2/c^4.

D.3 Fixing of q^2 for four-dimensional acceptance

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0025	F_L	0.0088	F_L	0.0089	F_L	0.0089	F_L	0.0043
S_3	0.0014	S_3	0.0005	S_3	0.0002	S_3	0.0005	S_3	0.0003
S_4	0.0037	S_4	0.0029	S_4	0.0002	S_4	0.0009	S_4	0.0001
S_5	0.0014	S_5	0.0005	S_5	0.0019	S_5	0.0020	S_5	0.0004
A_{FB}	0.0028	A_{FB}	0.0043	A_{FB}	0.0034	A_{FB}	0.0013	A_{FB}	0.0020
S_7	0.0001	S_7	0.0001	S_7	0.0002	S_7	0.0002	S_7	0.0001
S_8	0.0003	S_8	0.0001	S_8	0.0002	S_8	0.0000	S_8	0.0001
S_9	0.0000	S_9	0.0000	S_9	0.0001	S_9	0.0001	S_9	0.0001

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0034	F_L	0.0226	F_L	0.0013
S_3	0.0033	S_3	0.0102	S_3	0.0009
S_4	0.0030	S_4	0.0062	S_4	0.0005
S_5	0.0024	S_5	0.0044	S_5	0.0005
A_{FB}	0.0014	A_{FB}	0.0063	A_{FB}	0.0009
S_7	0.0002	S_7	0.0000	S_7	0.0001
S_8	0.0002	S_8	0.0000	S_8	0.0002
S_9	0.0001	S_9	0.0000	S_9	0.0000

Table D.7: Systematic uncertainties from fixing q^2 of the four-dimensional acceptance. Ranges of q^2 bins are given in GeV^2/c^4.

D.4 Higher order acceptance model

parameter	higher order result	parameter	nominal result
S_{1s}	0.330 ± 0.001	S_1^s	0.331 ± 0.001
S_3	0.001 ± 0.002	S_3	0.000 ± 0.002
S_4	-0.274 ± 0.002	S_4	-0.276 ± 0.002
S_5	-0.005 ± 0.002	S_5	-0.002 ± 0.002
S_{6s}	0.002 ± 0.002	S_6^s	0.002 ± 0.002
S_7	0.001 ± 0.002	S_7	0.001 ± 0.002
S_8	-0.050 ± 0.002	S_8	-0.050 ± 0.002
S_9	-0.085 ± 0.002	S_9	-0.087 ± 0.002
F_S	0.073 ± 0.003	F_S	0.083 ± 0.003
S_{S1}	-0.235 ± 0.003	S_{S1}	-0.229 ± 0.003
S_{S2}	0.002 ± 0.002	S_{S2}	0.001 ± 0.002
S_{S3}	0.001 ± 0.002	S_{S3}	0.003 ± 0.002
S_{S4}	0.001 ± 0.002	S_{S4}	0.001 ± 0.002
S_{S5}	-0.065 ± 0.002	S_{S5}	-0.065 ± 0.002

Table D.8: Results from the angular fit of the control decay $B^0 \to J/\psi K^{*0}$ using (left) the higher order acceptance description detailed in Section. 5.1.1.4 and (right) the nominal acceptance correction.

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0037	F_L	0.0108	F_L	0.0022	F_L	0.0033	F_L	0.0034
S_3	0.0007	S_3	0.0015	S_3	0.0003	S_3	0.0003	S_3	0.0010
S_4	0.0042	S_4	0.0007	S_4	0.0017	S_4	0.0015	S_4	0.0014
S_5	0.0162	S_5	0.0065	S_5	0.0025	S_5	0.0020	S_5	0.0021
A_{FB}	0.0004	A_{FB}	0.0020	A_{FB}	0.0000	A_{FB}	0.0011	A_{FB}	0.0021
S_7	0.0036	S_7	0.0030	S_7	0.0005	S_7	0.0001	S_7	0.0009
S_8	0.0003	S_8	0.0030	S_8	0.0009	S_8	0.0006	S_8	0.0013
S_9	0.0017	S_9	0.0013	S_9	0.0012	S_9	0.0008	S_9	0.0020

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0018	F_L	0.0084	F_L	0.0032	F_L	0.0001	F_L	0.0003
S_3	0.0013	S_3	0.0025	S_3	0.0005	S_3	0.0000	S_3	0.0010
S_4	0.0009	S_4	0.0002	S_4	0.0003	S_4	0.0006	S_4	0.0002
S_5	0.0022	S_5	0.0031	S_5	0.0028	S_5	0.0001	S_5	0.0002
A_{FB}	0.0004	A_{FB}	0.0059	A_{FB}	0.0018	A_{FB}	0.0001	A_{FB}	0.0004
S_7	0.0016	S_7	0.0112	S_7	0.0014	S_7	0.0014	S_7	0.0050
S_8	0.0018	S_8	0.0027	S_8	0.0023	S_8	0.0011	S_8	0.0003
S_9	0.0024	S_9	0.0015	S_9	0.0008	S_9	0.0007	S_9	0.0002

Table D.9: The effect of using a the nominal instead of a higher order acceptance model.

D.5 Peaking backgrounds

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0047	F_L	0.0018	F_L	0.0006	F_L	0.0006	F_L	0.0001
S_3	0.0023	S_3	0.0008	S_3	0.0010	S_3	0.0036	S_3	0.0100
S_4	0.0028	S_4	0.0012	S_4	0.0037	S_4	0.0034	S_4	0.0019
S_5	0.0028	S_5	0.0032	S_5	0.0007	S_5	0.0018	S_5	0.0005
A_{FB}	0.0062	A_{FB}	0.0075	A_{FB}	0.0018	A_{FB}	0.0001	A_{FB}	0.0022
S_7	0.0038	S_7	0.0002	S_7	0.0050	S_7	0.0031	S_7	0.0001
S_8	0.0020	S_8	0.0034	S_8	0.0048	S_8	0.0019	S_8	0.0057
S_9	0.0030	S_9	0.0024	S_9	0.0060	S_9	0.0014	S_9	0.0016

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0030	F_L	0.0020	F_L	0.0018	F_L	0.0055	F_L	0.0046
S_3	0.0018	S_3	0.0029	S_3	0.0000	S_3	0.0032	S_3	0.0050
S_4	0.0018	S_4	0.0029	S_4	0.0080	S_4	0.0012	S_4	0.0037
S_5	0.0014	S_5	0.0019	S_5	0.0058	S_5	0.0041	S_5	0.0032
A_{FB}	0.0014	A_{FB}	0.0032	A_{FB}	0.0013	A_{FB}	0.0058	A_{FB}	0.0056
S_7	0.0050	S_7	0.0037	S_7	0.0047	S_7	0.0005	S_7	0.0025
S_8	0.0013	S_8	0.0043	S_8	0.0034	S_8	0.0042	S_8	0.0014
S_9	0.0009	S_9	0.0004	S_9	0.0051	S_9	0.0021	S_9	0.0012

Table D.10: Deviations from the nominal observables due to the peaking backgrounds. The background events have been sampled from the corresponding charmonium mode.

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0064	F_L	0.0035	F_L	0.0024	F_L	0.0041	F_L	0.0023
S_3	0.0015	S_3	0.0029	S_3	0.0048	S_3	0.0050	S_3	0.0064
S_4	0.0039	S_4	0.0001	S_4	0.0052	S_4	0.0069	S_4	0.0097
S_5	0.0040	S_5	0.0040	S_5	0.0051	S_5	0.0085	S_5	0.0082
A_{FB}	0.0031	A_{FB}	0.0005	A_{FB}	0.0033	A_{FB}	0.0013	A_{FB}	0.0040
S_7	0.0011	S_7	0.0011	S_7	0.0017	S_7	0.0037	S_7	0.0009
S_8	0.0066	S_8	0.0005	S_8	0.0032	S_8	0.0032	S_8	0.0026
S_9	0.0014	S_9	0.0042	S_9	0.0052	S_9	0.0009	S_9	0.0011

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0034	F_L	0.0037	F_L	0.0017	F_L	0.0063	F_L	0.0027
S_3	0.0015	S_3	0.0049	S_3	0.0031	S_3	0.0005	S_3	0.0038
S_4	0.0049	S_4	0.0061	S_4	0.0047	S_4	0.0009	S_4	0.0020
S_5	0.0068	S_5	0.0038	S_5	0.0033	S_5	0.0010	S_5	0.0008
A_{FB}	0.0054	A_{FB}	0.0074	A_{FB}	0.0022	A_{FB}	0.0034	A_{FB}	0.0055
S_7	0.0014	S_7	0.0010	S_7	0.0016	S_7	0.0055	S_7	0.0002
S_8	0.0025	S_8	0.0033	S_8	0.0042	S_8	0.0066	S_8	0.0003
S_9	0.0034	S_9	0.0033	S_9	0.0032	S_9	0.0037	S_9	0.0012

Table D.11: Deviations from the nominal observables due to the peaking backgrounds. The background events have been generated using the kernel method described in the text.

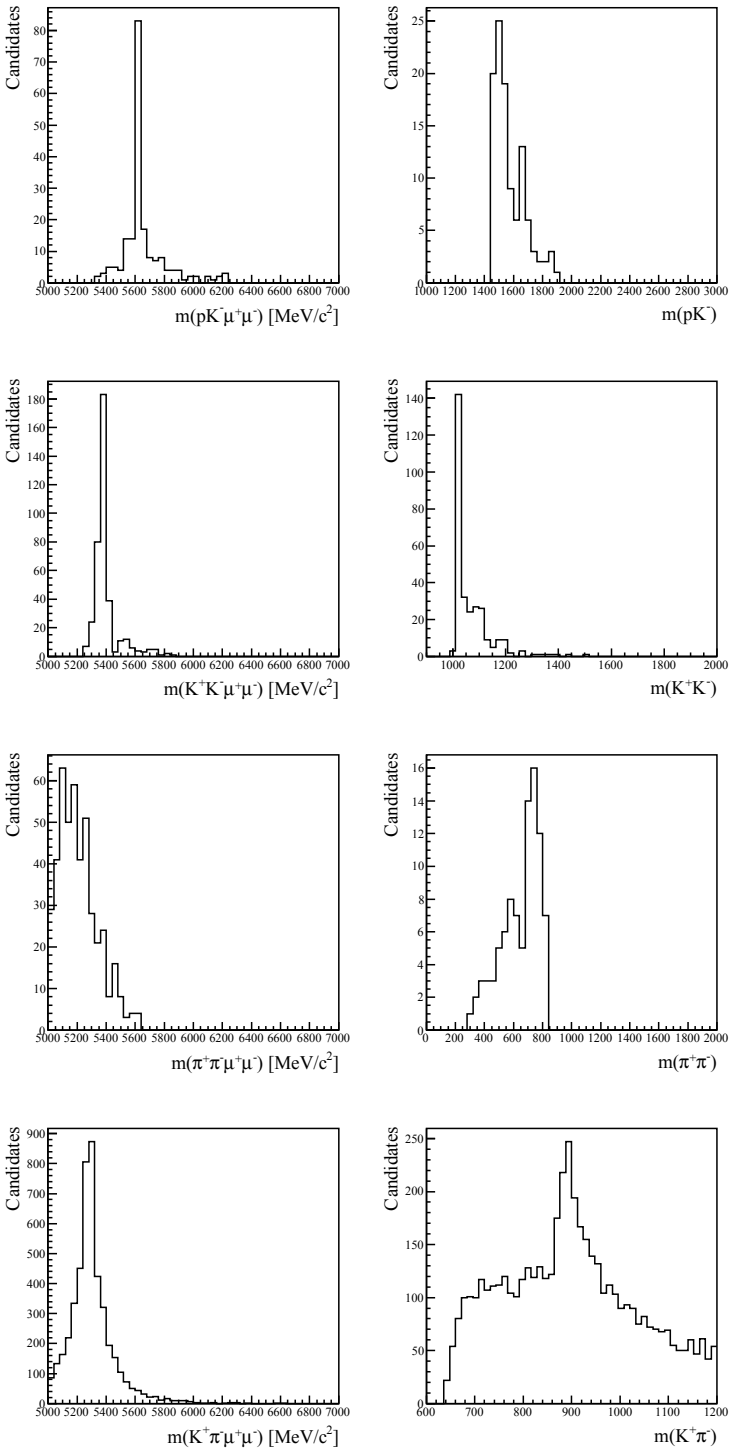


Figure D.1: Selected (first row) $\Lambda_b^0 \to pK\mu^+\mu^-$, (second row) $B_s^0 \to \phi\mu^+\mu^-$ and (third row) $B^0 \to \pi^+\pi^-\mu^+\mu^-$ peaking background events, as well as (fourth row) $B^0 \to K^{*0}\mu^+\mu^-$ swaps. The left column gives the reconstructed mass of the b hadron, the right column the reconstructed mass of the final state hadron.

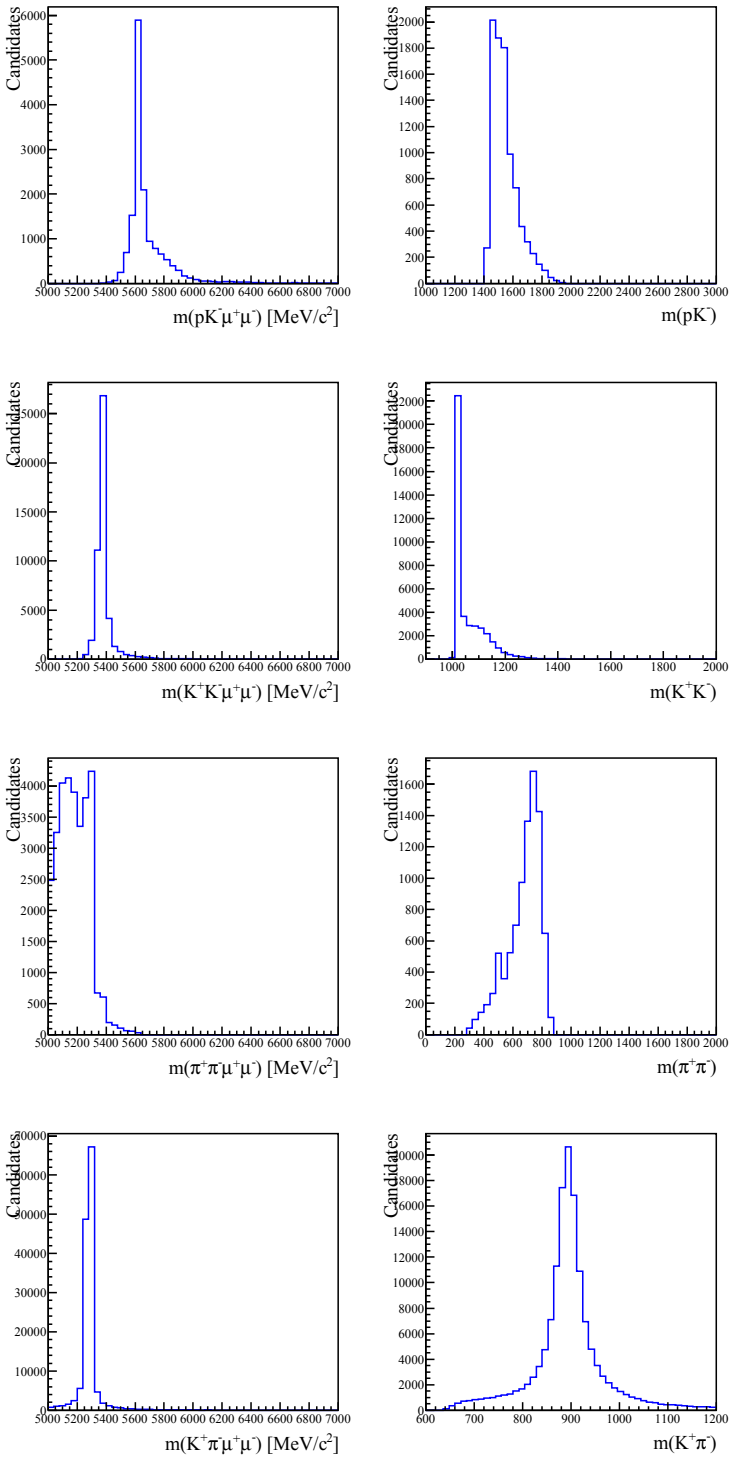


Figure D.2: Selected (first row) $\Lambda_b^0 \to J/\psi\, pK$, (second row) $B_s^0 \to J/\psi\, \phi$ and (third row) $B^0 \to J/\psi\, \pi^+\pi^-$ peaking background events, as well as (fourth row) $B^0 \to J/\psi\, K^{*0}$ swaps. The left column gives the reconstructed mass of the b hadron, the right column the reconstructed mass of the final state hadron.

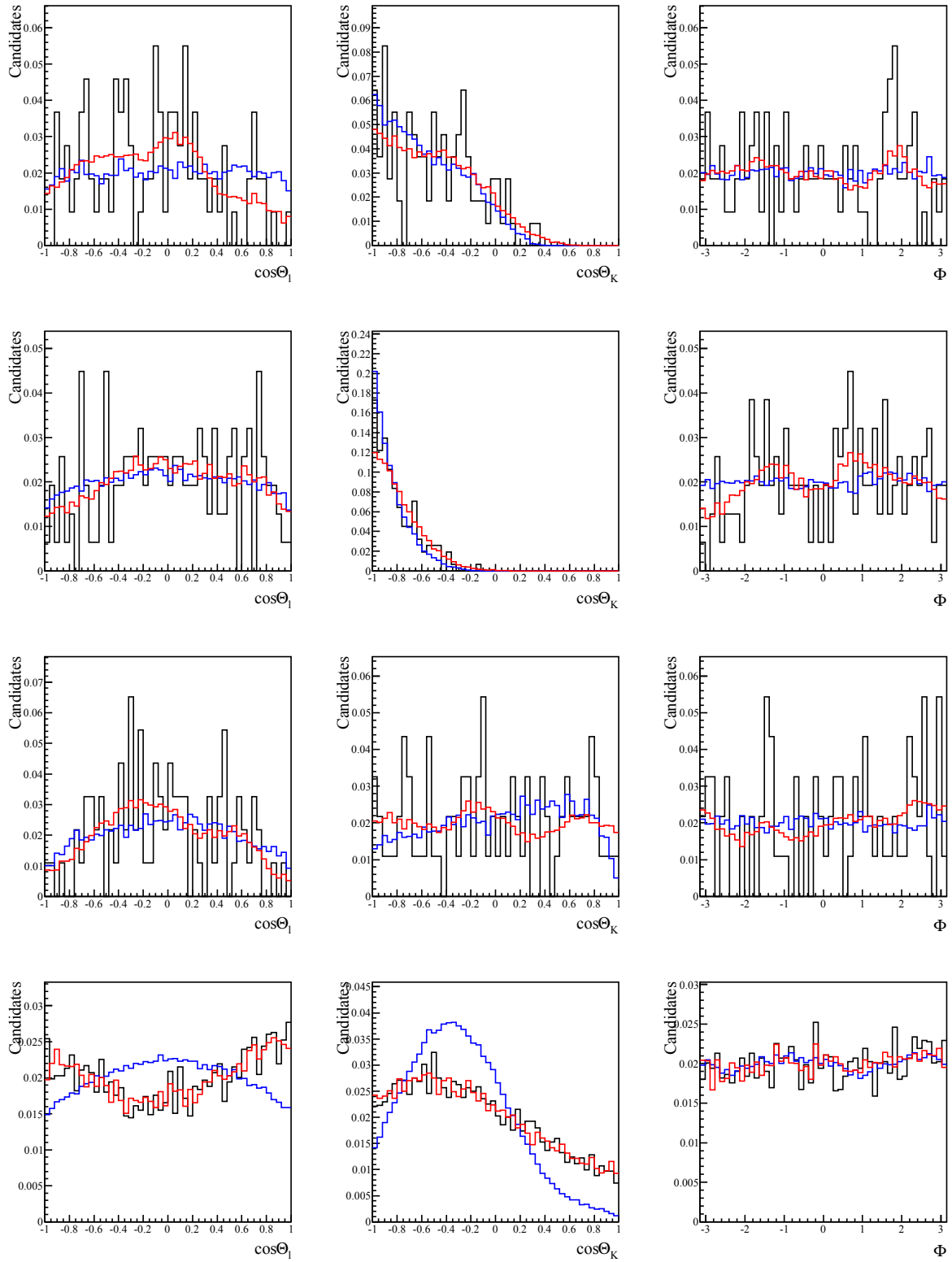


Figure D.3: The decay angles (left) $\cos\theta_l$, (middle) $\cos\theta_k$, and (right) ϕ for (black) $b \to X\mu^+\mu^-$ decays, (blue) $b \to J/\psi X$ decays, and (red) from the kernel method described in the text. The three peaking backgrounds studied are (first row) $\Lambda_b^0 \to pK\mu^+\mu^-$, (second row) $B_s^0 \to \phi\mu^+\mu^-$, (third row) $B^0 \to \pi^+\pi^-\mu^+\mu^-$ and (fourth row) $B^0 \to K^+\pi^-\mu^+\mu^-$ swaps.

D.6 Angular background modeling

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0003	F_L	0.0033	F_L	0.0013	F_L	0.0010	F_L	0.0012
S_3	0.0010	S_3	0.0031	S_3	0.0024	S_3	0.0022	S_3	0.0039
S_4	0.0007	S_4	0.0004	S_4	0.0003	S_4	0.0010	S_4	0.0016
S_5	0.0002	S_5	0.0010	S_5	0.0010	S_5	0.0010	S_5	0.0034
A_{FB}	0.0001	A_{FB}	0.0009	A_{FB}	0.0014	A_{FB}	0.0005	A_{FB}	0.0008
S_7	0.0001	S_7	0.0005	S_7	0.0002	S_7	0.0002	S_7	0.0003
S_8	0.0000	S_8	0.0004	S_8	0.0002	S_8	0.0002	S_8	0.0001
S_9	0.0006	S_9	0.0013	S_9	0.0019	S_9	0.0033	S_9	0.0018

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0011	F_L	0.0003	F_L	0.0007	F_L	0.0021	F_L	0.0005
S_3	0.0002	S_3	0.0059	S_3	0.0011	S_3	0.0010	S_3	0.0008
S_4	0.0012	S_4	0.0002	S_4	0.0014	S_4	0.0017	S_4	0.0015
S_5	0.0005	S_5	0.0001	S_5	0.0024	S_5	0.0005	S_5	0.0010
A_{FB}	0.0002	A_{FB}	0.0002	A_{FB}	0.0021	A_{FB}	0.0013	A_{FB}	0.0000
S_7	0.0000	S_7	0.0001	S_7	0.0003	S_7	0.0002	S_7	0.0001
S_8	0.0000	S_8	0.0002	S_8	0.0002	S_8	0.0001	S_8	0.0000
S_9	0.0004	S_9	0.0018	S_9	0.0034	S_9	0.0004	S_9	0.0007

Table D.12: Systematic effect due to the angular background modeling.

$0.1 < q^2 < 0.98\,\mathrm{GeV}^2/c^4$:

$1.1 < q^2 < 2.5\,\mathrm{GeV}^2/c^4$:

$2.5 < q^2 < 4.0\,\mathrm{GeV}^2/c^4$:

$4.0 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$:

$6.0 < q^2 < 8.0\,\mathrm{GeV}^2/c^4$:

$15.0 < q^2 < 17.0\,\mathrm{GeV}^2/c^4$:

$17.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$:

$11.0 < q^2 < 12.5\,\mathrm{GeV}^2/c^4$:

$1.1 < q^2 < 6.0\,\mathrm{GeV}^2/c^4$:

$15.0 < q^2 < 19.0\,\mathrm{GeV}^2/c^4$:

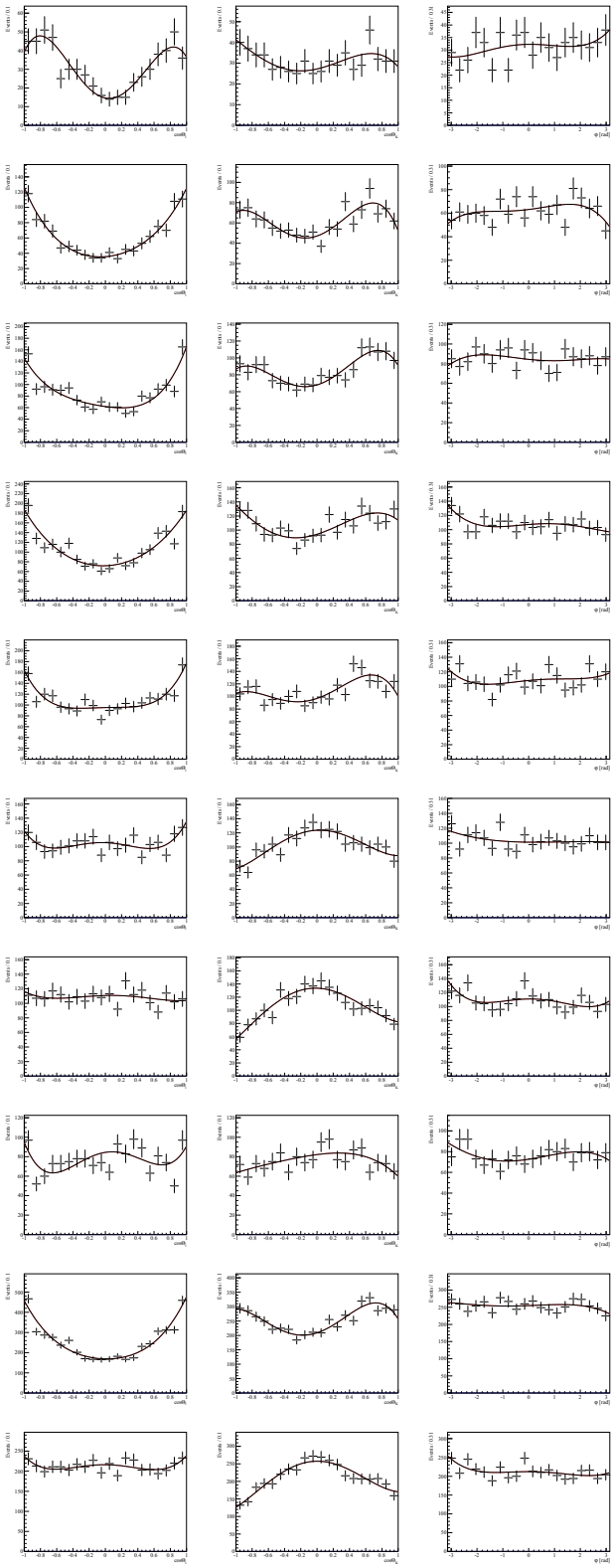


Figure D.4: The angular distribution of combinatorial background events in the high mass sideband ($m_{K\pi\mu\mu} \in [5355, 5700]\,\mathrm{MeV}/c^2$).

D.7 Signal mass modeling

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0009	F_L	0.0021	F_L	0.0022	F_L	0.0013	F_L	0.0006
S_3	0.0001	S_3	0.0001	S_3	0.0003	S_3	0.0004	S_3	0.0006
S_4	0.0000	S_4	0.0001	S_4	0.0003	S_4	0.0004	S_4	0.0006
S_5	0.0008	S_5	0.0006	S_5	0.0007	S_5	0.0013	S_5	0.0019
A_{FB}	0.0005	A_{FB}	0.0011	A_{FB}	0.0003	A_{FB}	0.0007	A_{FB}	0.0011
S_7	0.0000	S_7	0.0005	S_7	0.0001	S_7	0.0001	S_7	0.0002
S_8	0.0000	S_8	0.0001	S_8	0.0000	S_8	0.0001	S_8	0.0002
S_9	0.0000	S_9	0.0001	S_9	0.0002	S_9	0.0004	S_9	0.0002

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0013	F_L	0.0022	F_L	0.0011	F_L	0.0017	F_L	0.0017
S_3	0.0039	S_3	0.0072	S_3	0.0014	S_3	0.0001	S_3	0.0037
S_4	0.0003	S_4	0.0011	S_4	0.0005	S_4	0.0003	S_4	0.0005
S_5	0.0019	S_5	0.0020	S_5	0.0019	S_5	0.0005	S_5	0.0019
A_{FB}	0.0023	A_{FB}	0.0034	A_{FB}	0.0019	A_{FB}	0.0001	A_{FB}	0.0027
S_7	0.0000	S_7	0.0003	S_7	0.0000	S_7	0.0003	S_7	0.0000
S_8	0.0001	S_8	0.0000	S_8	0.0000	S_8	0.0002	S_8	0.0000
S_9	0.0001	S_9	0.0001	S_9	0.0000	S_9	0.0001	S_9	0.0001

Table D.13: Systematic effect of the signal mass model.

D.8 $m_{K\pi}$ related systematic uncertainties

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0002	F_L	0.0004	F_L	0.0004	F_L	0.0002	F_L	0.0001
S_3	0.0000	S_3	0.0000	S_3	0.0000	S_3	0.0000	S_3	0.0000
S_4	0.0001	S_4	0.0000	S_4	0.0001	S_4	0.0001	S_4	0.0002
S_5	0.0004	S_5	0.0002	S_5	0.0002	S_5	0.0004	S_5	0.0004
A_{FB}	0.0001	A_{FB}	0.0001	A_{FB}	0.0000	A_{FB}	0.0000	A_{FB}	0.0001
S_7	0.0000	S_7	0.0001	S_7	0.0000	S_7	0.0000	S_7	0.0000
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0003	F_L	0.0010	F_L	0.0001	F_L	0.0003	F_L	0.0003
S_3	0.0007	S_3	0.0035	S_3	0.0001	S_3	0.0000	S_3	0.0011
S_4	0.0006	S_4	0.0024	S_4	0.0002	S_4	0.0001	S_4	0.0009
S_5	0.0010	S_5	0.0026	S_5	0.0005	S_5	0.0002	S_5	0.0013
A_{FB}	0.0008	A_{FB}	0.0024	A_{FB}	0.0003	A_{FB}	0.0000	A_{FB}	0.0011
S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0000
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

Table D.14: Systematic deviations due to using the ISOBAR model instead of the nominal LASS description of the S-wave.

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0004	F_L	0.0024	F_L	0.0050	F_L	0.0035	F_L	0.0016
S_3	0.0000	S_3	0.0000	S_3	0.0002	S_3	0.0003	S_3	0.0005
S_4	0.0003	S_4	0.0002	S_4	0.0011	S_4	0.0020	S_4	0.0019
S_5	0.0009	S_5	0.0011	S_5	0.0024	S_5	0.0047	S_5	0.0047
A_{FB}	0.0003	A_{FB}	0.0009	A_{FB}	0.0006	A_{FB}	0.0006	A_{FB}	0.0014
S_7	0.0001	S_7	0.0003	S_7	0.0005	S_7	0.0004	S_7	0.0003
S_8	0.0000	S_8	0.0001	S_8	0.0001	S_8	0.0001	S_8	0.0001
S_9	0.0000	S_9	0.0000	S_9	0.0001	S_9	0.0000	S_9	0.0000

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0006	F_L	0.0004	F_L	0.0006	F_L	0.0042	F_L	0.0013
S_3	0.0013	S_3	0.0015	S_3	0.0012	S_3	0.0001	S_3	0.0042
S_4	0.0012	S_4	0.0010	S_4	0.0021	S_4	0.0013	S_4	0.0035
S_5	0.0018	S_5	0.0011	S_5	0.0044	S_5	0.0026	S_5	0.0048
A_{FB}	0.0015	A_{FB}	0.0010	A_{FB}	0.0027	A_{FB}	0.0004	A_{FB}	0.0042
S_7	0.0000	S_7	0.0000	S_7	0.0000	S_7	0.0006	S_7	0.0000
S_8	0.0000	S_8	0.0000	S_8	0.0000	S_8	0.0001	S_8	0.0000
S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000	S_9	0.0000

Table D.15: Systematic uncertainties due to the $m_{K\pi}$ parametrisation of background events.

$0.1 < q^2 < 1.0$		$1.0 < q^2 < 2.5$		$2.5 < q^2 < 4.0$		$4.0 < q^2 < 6.0$		$6.0 < q^2 < 8.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0007	F_L	0.0022	F_L	0.0010	F_L	0.0033	F_L	0.0031
S_3	0.0008	S_3	0.0006	S_3	0.0034	S_3	0.0028	S_3	0.0016
S_4	0.0025	S_4	0.0003	S_4	0.0023	S_4	0.0002	S_4	0.0030
S_5	0.0011	S_5	0.0029	S_5	0.0025	S_5	0.0012	S_5	0.0054
A_{FB}	0.0005	A_{FB}	0.0006	A_{FB}	0.0031	A_{FB}	0.0012	A_{FB}	0.0019
S_7	0.0019	S_7	0.0009	S_7	0.0020	S_7	0.0023	S_7	0.0023
S_8	0.0033	S_8	0.0008	S_8	0.0033	S_8	0.0031	S_8	0.0018
S_9	0.0009	S_9	0.0003	S_9	0.0025	S_9	0.0016	S_9	0.0034

$15.0 < q^2 < 17.0$		$17.0 < q^2 < 19.0$		$11.0 < q^2 < 12.5$		$1.1 < q^2 < 6.0$		$15.0 < q^2 < 19.0$	
param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$	param.	$\sigma_{\text{syst.}}$
F_L	0.0020	F_L	0.0008	F_L	0.0006	F_L	0.0001	F_L	0.0016
S_3	0.0017	S_3	0.0015	S_3	0.0013	S_3	0.0017	S_3	0.0034
S_4	0.0009	S_4	0.0014	S_4	0.0013	S_4	0.0027	S_4	0.0036
S_5	0.0012	S_5	0.0009	S_5	0.0023	S_5	0.0002	S_5	0.0060
A_{FB}	0.0007	A_{FB}	0.0013	A_{FB}	0.0005	A_{FB}	0.0018	A_{FB}	0.0036
S_7	0.0010	S_7	0.0056	S_7	0.0017	S_7	0.0007	S_7	0.0031
S_8	0.0010	S_8	0.0007	S_8	0.0020	S_8	0.0041	S_8	0.0026
S_9	0.0013	S_9	0.0020	S_9	0.0019	S_9	0.0021	S_9	0.0008

Table D.16: Systematic uncertainties due to the assumption of flat efficiency in $m_{K\pi}$.

Appendix E

Input variable correlation

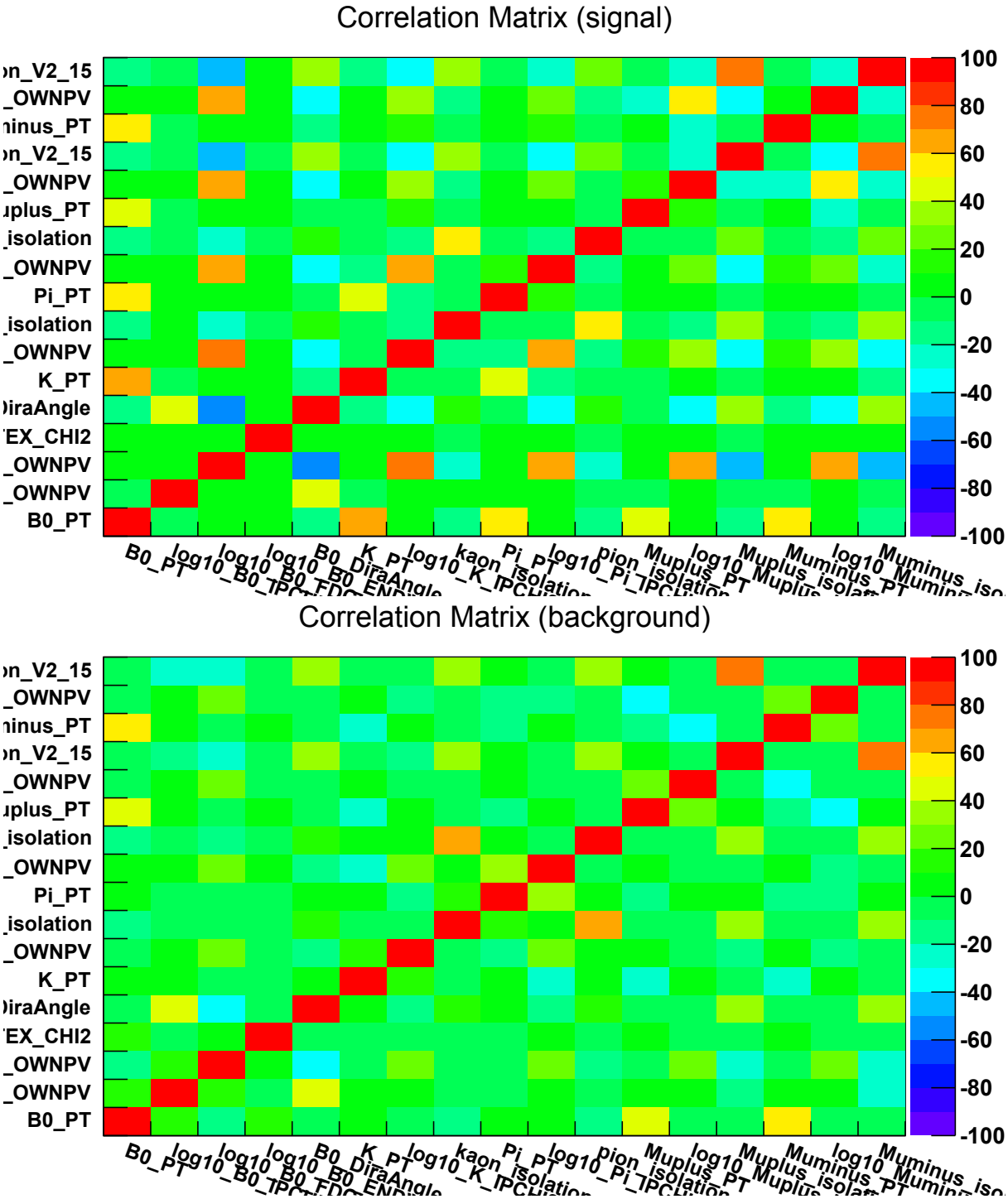


Figure E.1: Correlation matrix for the signal (top) and background (bottom) samples used for the BDT classifier on the R_{K^*} analysis.

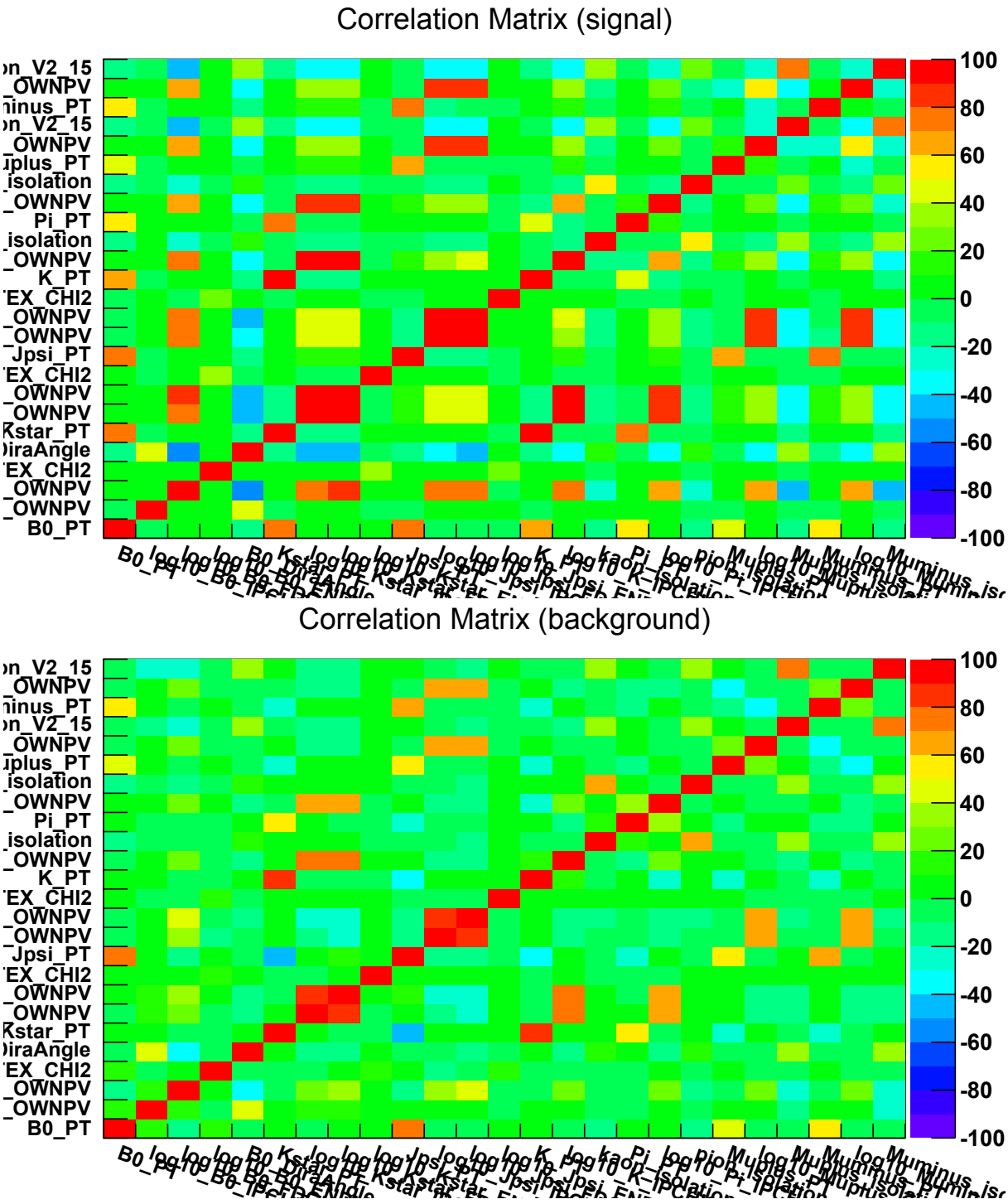


Figure E.2: Correlation matrix for the signal (top) and background (bottom) samples tested based on the classifier used in the $B^0 \to K^{*0} e^+ e^-$ analysis [**?**] where information from K^{*0} and $\ell^+ \ell^-$ were used.